

STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 109965

TO: Ruixiang Li
Location: CM1/10D19/10E18
Art Unit: 1646
Tuesday, December 16, 2003

Case Serial Number: 09/900448

From: Barb O'Bryen
Location: Biotech-Chem Library
CM1-6A05
Phone: 308-4291 *BOB*

barbara.obryen@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 23:41:46 ; Search time 25 Seconds

(without alignments)
735,498 Million cell updates/sec

Title: US-09-900-448-2

Perfect score: 2185
Sequence: 1 MARVLGAPVAGLWMSLCSL.....NAKALPQPNVTSLLGCTH 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2139.5	97.9	462	1	HEMO_HUMAN
2	1700.5	77.8	460	1	HEMO_RABIT
3	1556.5	71.2	460	1	HEMO_RAT
4	1545.5	70.7	460	1	HEMO_MOUSE
5	1544	70.7	459	1	HEMO_PIG
6	203	9.3	578	1	MM17_MOUSE
7	202.5	9.3	470	1	MM12_HUMAN
8	199	9.1	606	1	MM12_RABIT
9	196	9.0	464	1	MM12_MOUSE
10	195.5	8.9	476	1	MM10_MOUSE
11	195.5	8.9	477	1	MM03_HORSE
12	189.5	8.7	476	1	MM10_HUMAN
13	187	8.6	476	1	MM10_RAT
14	184.5	8.4	468	1	MM01_RABIT
15	184.5	8.4	492	1	MM01_MOUSE
16	183.5	8.4	467	1	MM08_HUMAN
17	181.5	8.3	469	1	MM01_HUMAN
18	179.5	8.2	488	1	MM11_HUMAN
19	177.5	8.1	471	1	MM13_RABIT
20	177	8.1	475	1	MM03_RAT
21	177	8.1	477	1	MM03_HUMAN
22	176	8.1	471	1	MM13_HUMAN
23	176	8.1	472	1	MM13_MOUSE
24	174.5	8.0	469	1	MM01_HORSE
25	173	7.9	477	1	MM03_MOUSE
26	172	7.9	469	1	MM15_HUMAN
27	169.5	7.8	477	1	MM11_XENLA
28	169.5	7.8	478	1	VTNC_HUMAN
29	167.5	7.7	508	1	MM19_HUMAN
30	166.5	7.6	466	1	MM13_RAT
31	166	7.6	469	1	MM01_PIG
32	166	7.6	478	1	MM03_RABIT
33	165.5	7.6	478	1	VTNC_MOUSE

34	165	7.6	657	1	MM15_MOUSE	O54732 mus musculus
35	163	7.5	618	1	MM24_MOUSE	O90962 mus musculus
36	163	7.5	618	1	MM24_RAT	O99962 rattus norv
37	162.5	7.4	645	1	MM24_HUMAN	O99962 mus musculus
38	161.5	7.4	465	1	MM08_MOUSE	O87038 mus musculus
39	159.5	7.3	466	1	MM08_RAT	O87038 rattus norv
40	158	7.2	459	1	VTNC_PIG	O48819 sus scrofa
41	155.5	7.1	467	1	MM18_XENLA	O13065 xenopus lae
42	154.5	7.1	472	1	MM13_MOUSE	P33435 mus musculus
43	152	7.0	607	1	MM16_HUMAN	P51512 bos taurus
44	151	6.9	469	1	MM01_BOVIN	P28053 bos taurus
45	151	6.9	607	1	MM16_MOUSE	O9wrt0 mus musculus

ALIGNMENTS

RESULT 1	ID	HEMO_HUMAN	STANDARD;	PRT;	462 AA.
AC	P02790;	21-JUL-1986 (Rel. 01, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Hemopexin precursor (Beta-1B-glycoprotein).				
CN	HPX.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88316972; PubMed=2842511; Silengo L.;				
RA	Altuda F., Poli V., Restagno G., Silengo L.;				
RT	"Structure of the human hemopexin gene and evidence for				
RT	intron-mediated evolution."				
RL	J. Mol. Evol. 27:102-108(1988).				
RN	[2]				
RP	SEQUENCE OF 2-462 FROM N.A.				
RX	MEDLINE=89122012; PubMed=3220477;				
RA	Law M.L., Cai G.Y., Hartz J.A., Jones C., Kao F.T.;				
RT	"The hemopexin gene maps to the same location as the beta-globin gene				
RL	cluster on human chromosome 11."				
RN	Genomics 3:48-52(1988).				
RP	SEQUENCE OF 22-462 FROM N.A.				
RX	MEDLINE=85242073; PubMed=2989777;				
RA	Altuda F., Poli V., Restagno G., Argos P., Cortese R., Silengo L.;				
RT	"The primary structure of human hemopexin deduced from cDNA sequence:				
RT	evidence for internal, repeating homology."				
RL	Nucleic Acids Res. 13:3841-3859(1985).				
RN	[4]				
RP	SEQUENCE OF ACTIVE PROTEIN.				
RX	MEDLINE=85113173; PubMed=3855550;				
RA	Takahashi N., Takahashi Y., Putnam F.W.;				
RT	"Complete amino acid sequence of human hemopexin, the heme-binding				
RT	protein of serum."				
RL	Proc. Natl. Acad. Sci. U.S.A. 82:73-77(1985).				
RN	[5]				
RP	SEQUENCE OF 24-255.				
RX	MEDLINE=85076955; PubMed=6510521;				
RA	Frantkova V., Borvak J., Kluh I., Moravsek L.;				
RT	"Amino acid sequence of the N-terminal region of human hemopexin."				
RL	FEBS Lett. 178:213-216(1984).				
RN	[6]				
RP	PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.				
RX	MEDLINE=84193947; PubMed=6371807;				
RA	Takahashi N., Takahashi Y., Putnam F.W.;				
RT	"Structure of human hemopexin: O-glycosyl and N-glycosyl sites and				
RT	unusual clustering of tryptophan residues."				
RL	Proc. Natl. Acad. Sci. U.S.A. 81:2021-2025(1984).				
CC	-1- FUNCTION: BINDS HEME AND TRANSPORTS IT TO THE LIVER FOR BREAKDOWN				
	AND IRON RECOVERY, AFTER WHICH THE FREE HEMOPEXIN RETURNS TO THE				

CIRCULATION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
 CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC
 DR EMBL: M36803; AA58678.1; -
 DR EMBL: M36796; AA58678.1; JOINED.
 DR EMBL: M36799; AA58678.1; JOINED.
 DR EMBL: M36800; AA58678.1; JOINED.
 DR EMBL: M36801; AA58678.1; JOINED.
 DR EMBL: M36802; AA58678.1; JOINED.
 DR EMBL: J03048; AA52704.1; -
 DR EMBL: X02537; CA26382.1; ALT_INIT.
 DR PIR: I56456; OOHU.
 DR HSSP: P20058; 1HXN.
 DR SWISS-2DPAGE: P02790; HUMAN.
 DR Siens-2DPAGE: P02790; -
 DR Genew: HGNC:5171; HPX.
 DR MIM: 142290; -
 DR GO: GO:0005615; C:extracellular space; TAS.
 DR GO: GO:0005488; F:binding activity; TAS.
 DR GO: GO:0015332; F:heme transporter activity; TAS.
 DR GO: GO:0005211; F:plasma glycoprotein; TAS.
 DR GO: GO:0015886; F:heme transporter; TAS.
 DR GO: GO:0006879; F:iron ion homeostasis; TAS.
 DR InterPro: IPR000585; Hemopexin.
 DR Pfam: PF00045; hemopexin; 5.
 DR SMART: SM00120; HX; 5.
 DR PROSITE: PS00024; HEMOPEXIN; 2.
 DR KMW Glycoprotein; Heme; Plasma; Repeat; Transport; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 462
 FT DOMAIN 56 93 HEMOPEXIN-LIKE 1.
 FT DOMAIN 97 141 HEMOPEXIN-LIKE 2.
 FT DOMAIN 188 231 HEMOPEXIN-LIKE 3.
 FT DOMAIN 263 306 HEMOPEXIN-LIKE 4.
 FT DOMAIN 308 351 HEMOPEXIN-LIKE 5.
 FT METAL 79 79 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL 150 150 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT DISULFID 50 231
 FT DISULFID 149 154
 FT DISULFID 188 200
 FT DISULFID 257 460
 FT DISULFID 366 408
 FT DISULFID 418 435
 FT CARBOHYD 24 24
 FT CARBOHYD 64 64 O-LINKED.
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. .).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. .).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. .).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. .).
 SQ SEQUENCE 462 AA; 51676 MW; 054844D060376388 CRC64;
 Query Match 97.9%; Score 2139.5; DB 1; Length 462;
 Best Local Similarity 84.6%; Pred. No. 5.8e-176;
 Matches 391; Conservative 0; Mismatches 0; Indels 71; Gaps 1;
 QY 1 MARVLGAPVALGLMSLCMSLAITPLPPTSAHGNVAGETKPPDPVTERGSDGMSFPDATT 60
 DB 1 MAVLGAPVALGLMSLCMSLAITPLPPTSAHGNVAGETKPPDPVTERGSDGMSFPDATT 60
 QY 61 LDDNGTMLFFKGEFVWKSFKDRELISERKWNPPSPVDAAFROGHNSVFLIKDQVWVVP 120
 DB 61 LDDNGTMLFFKGEFVWKSFKDRELISERKWNPPSPVDAAFROGHNSVFLIKDQVWVVP 120

QY 121 PEKKEGYPRKLODEPFGIPSPPLDAVCHRGCEQAGVLFQ----- 163
 DB 121 PEKKEGYPRKLODEPFGIPSPPLDAVCHRGCEQAGVLFQGDENWMTDATGTMKER 180
 QY 164 -----GHGRN 169
 DB 181 SMPAVGNCSSALRWLGRRYCFQGNQPLRFPDVRGVEPRPRDVRDYMPCPGRGHGRN 240
 QY 170 GTGHGNTGTHGPRYMCSPHLVLSALTSNDHGTVPFGTHWRRLDTSRGMSWPIAHQ 229
 DB 241 GTGHGNTGTHGPRYMCSPHLVLSALTSNDHGTVPFGTHWRRLDTSRGMSWPIAHQ 300
 QY 230 WPGSPAVDAAPFSEBKLTVVQGYVFLTKGTYLVGSPYRLEKEVGTPHGILLDSV 289
 DB 301 WPGSPAVDAAPFSEBKLTVVQGYVFLTKGTYLVGSPYRLEKEVGTPHGILLDSV 360
 QY 290 DAAFCPGSSRLHIMAGRRLWDLKSGAQTWELPWHEKYDGLCNEKSLGPNCSGA 349
 DB 361 DAAFCPGSSRLHIMAGRRLWDLKSGAQTWELPWHEKYDGLCNEKSLGPNCSGA 420
 QY 350 NGPGLYLHGPNIYCYSDVEKLNAAKALPOPOVNTSLGCTH 391
 DB 421 NGPGLYLHGPNIYCYSDVEKLNAAKALPOPOVNTSLGCTH 462
 RESULT 2
 ID HEMO_RABBIT STANDARD; PRT; 460 AA.
 AC P20058;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hemopexin precursor.
 GN HPX.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND HEME-BINDING SITES.
 RC TISSUE=Liver;
 RX MEDLINE=93203213; PubMed=7681064;
 RA Morgan W.T., Master P., Tatum F., Kao S.-M., Alam J., Smith A.;
 RT "Identification of the histidine residues of hemopexin that
 RT coordinate with heme-iron and of a receptor-binding region";
 RL J. Biol. Chem. 268:6256-6262(1993).
 RN [2]
 RP SEQUENCE OF 26-53.
 RX MEDLINE=88339942; PubMed=3421961;
 RA Wellner D., Cheng K.C., Mueller-Eberhard U.;
 RT "N-terminal amino acid sequences of the hemopexins from chicken, rat
 RT and rabbit";
 RL Biochem. Biophys. Res. Commun. 155:622-625(1988).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 242-460.
 RX MEDLINE=96173004; PubMed=8590016;
 RA Faber H.R., Groom C.R., Baker H.M., Morgan W.T., Smith A., Baker E.N.;
 RT "1.8-A crystal structure of the C-terminal domain of rabbit serum
 RT haemopexin";
 RL Structure 3:551-559(1995).
 CC -1- FUNCTION: BINDS HEME AND TRANSPORTS IT TO THE LIVER FOR BREAKDOWN
 CC AND IRON RECOVERY. AFTER WHICH THE FREE HEMOPEXIN RETURNS TO THE
 CC CIRCULATION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
 CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>

CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
 CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M62642; AAA4137.1; -;
 CC EMBL: X60006; CAA42621.1; -;
 CC PIR: A43079; OORT.
 CC HSSP: P20058; 1HXN.
 CC InterPro: IPR000585; Hemopexin.
 CC Pfam: PF00045; hemopexin; 5.
 CC SMART: SM00120; HX; 5.
 CC DR Glycoprotein; Heme; Plasma; Repeat; Transport; Signal.
 CC KM Glycoprotein; Heme; Plasma; Repeat; Transport; Signal.
 CC FT SIGNAL 1 23
 CC FT CHAIN 24 460
 CC FT DOMAIN 56 93 HEMOPEXIN-LIKE 1.
 CC FT DOMAIN 97 140 HEMOPEXIN-LIKE 2.
 CC FT DOMAIN 187 230 HEMOPEXIN-LIKE 3.
 CC FT DOMAIN 261 304 HEMOPEXIN-LIKE 4.
 CC FT DOMAIN 306 349 HEMOPEXIN-LIKE 5.
 CC FT METAL 79 79 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC FT METAL 149 149 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC FT DISULFID 50 230 BY SIMILARITY.
 CC FT DISULFID 148 153 BY SIMILARITY.
 CC FT DISULFID 187 199 BY SIMILARITY.
 CC FT DISULFID 255 458 BY SIMILARITY.
 CC FT DISULFID 364 406 BY SIMILARITY.
 CC FT DISULFID 416 433 BY SIMILARITY.
 CC FT CARBOHYD 38 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CONFLICT 38 38 N-> C (IN REF. 2).
 CC FT CONFLICT 49 50 HC-> KW (IN REF. 2).
 CC SQ SEQUENCE 460 AA; 51291 MW; A1D67D05B5BFD83 CRC64;
 CC
 CC Query Match 71.2%; Score 1556.5; DB 1; Length 460;
 CC Best Local Similarity 63.8%; Pred. No. 5.8e-126; Indels 79; Gaps 6;
 CC Matches 236; Conservative 35; Mismatches 54;
 CC
 CC 1 MARVLGAPVALGLMSLCWSLAIAIPLPTSAHGNVAEGE--TRPDPVTERCSGDSWSPDA 58
 CC 1 MARTVALNITLVILGLCWSLAVANPLP--AAHETVAKGEGTKPDSVIEHCSDAMSFDA 58
 CC 59 TTIIDNGTMLFFFGGEFVWVSKMDRELISERWNPSPVDAAFRQGHNSVFLIKDKVWY 118
 CC 59 TITMDHNGTMLFFFGGEFVWVSKMDRELISERWNPSPVDAAFRQGHNSVFLIKDKVWY 117
 CC 119 YPEKKEKGVKPLQDEBPGISPLDAAVCHGEGCOABGVLFPOGH-- 165
 CC 118 YPEKKEKGVKPLQDEBPGISPLDAAVCHGEGCOABGVLFPOGH-- 165
 CC 166 -----GH-- 167
 CC 178 ERSWPAVNGCTALRWLERVYFCQGNKFLRFPNVTGEVPRRYLDARDYISCPGKHGK 217
 CC 168 RNGTGHGNSYTHGPEYMRCSPLVLIALTSNKGATVAESGTHYRLDTSRDGWSWPI 226
 CC 238 LNRGTAGHNSYTH--PMHSSCNADPGIALSLSDRGATVAFSGSHYRLDSSRGHWSWPI 295
 CC 227 AHWMPGSPAVDAAFSGWBEKLYVVOGTQVYVFLTKGTYLVSGYPRKLEKVEVTPHGILL 286
 CC 296 AHWMPGSPAVDAAFSGWBEKLYVVOGTQVYVFLTKGTYLVSGYPRKLEKVEVTPHGILL 355
 CC 287 DSVDAAFICPGSSRLIMAGRLMWLDLXSGAQTATELPWPEKVDGALCKEKSIGSPS 346

DB 356 DRIIDAFSCPGSSKLYVTSGRLMWLDLXSGAQTATELPWPEKVDGALCKEKSIGSPS 415
 CC
 CC 347 CSANGPGLVIRGPNLYCYSDVEKMAAKALPPQVNTSLGCT 390
 CC 416 CSANGPGLVIRGPNLYCYSDVEKMAAKALPPQVNTSLGCT 459
 CC
 CC RESULT 4
 CC HEMO MOUSE STANDARD; PRT; 460 AA.
 CC ID HEMO_MOUSE
 CC AC Q91X72; P97824; Q8WUP0;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Hemopexin precursor.
 CC GN HPX OR HPXN.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RP TISSUE=Liver;
 CC RC MEDLINE=22388257; PubMed=12477932;
 CC RA Klausner R.D., Collins F.S., Wagner L.H., Derge J.G.,
 CC Altschul S.F., Zeeberg B., Buicow K.H., Schaefer C.F., Bhut N.K.,
 CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
 CC Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
 CC Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
 CC Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Pange C.,
 CC Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,
 CC Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
 CC Valhalla D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 CC Fahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 CC Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 CC Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 CC Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 CC Roderfeld Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 CC Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 CC "Generation and initial analysis of more than 15,000 full-length
 CC human and mouse cDNA sequences.";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC
 CC [2]
 CC SEQUENCE OF 7-460 FROM N.A.
 CC RP TISSUE=Liver;
 CC RC Koepel R.R., Rohrbach D.H., Breckeliser B.B.;
 CC RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: BINDS HEME AND TRANSPORTS IT TO THE LIVER FOR BREAKDOWN
 CC AND IRON RECOVERY, AFTER WHICH THE FREE HEMOPEXIN RETURNS TO THE
 CC CIRCULATION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
 CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: BC011246; AAH11246.1; -;
 CC EMBL: BC019901; AAH19901.1; -;
 CC EMBL: U89889; AAB49490.1; -;
 CC MGD: MGJ:105112; HPXN.
 CC InterPro: IPR000585; Hemopexin.
 CC Pfam: PF00045; hemopexin; 5.
 CC SMART: SM00120; HX; 5.
 CC PROSITE: PS00024; HEMOPEXIN; 1.

KW Glycoprotein; Heme; Plasma; Repeat; Transport; Signal.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 460 HEMOPEXIN.

FT DOMAIN 56 93 HEMOPEXIN-LIKE 1.

FT DOMAIN 97 140 HEMOPEXIN-LIKE 2.

FT DOMAIN 187 230 HEMOPEXIN-LIKE 3.

FT DOMAIN 261 304 HEMOPEXIN-LIKE 4.

FT DOMAIN 306 349 HEMOPEXIN-LIKE 5.

FT METAL 79 79 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

FT METAL 149 149 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

FT DISULFID 50 230 BY SIMILARITY.

FT DISULFID 148 153 BY SIMILARITY.

FT DISULFID 187 199 BY SIMILARITY.

FT DISULFID 255 458 BY SIMILARITY.

FT DISULFID 364 406 BY SIMILARITY.

FT DISULFID 416 433 BY SIMILARITY.

FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 30 30 H -> N (IN REF. 2).

FT CONFLICT 233 233 R -> Q (IN REF. 2).

FT CONFLICT 453 453 MISSING (IN REF. 2).

SEQUENCE 460 AA; 51340 MM; 363BABCT520D1B39 CRC64;

Query Match 70.7%; Score 1545.5; DB 1; Length 460;

Best Local Similarity 62.9%; Pred. No. 5e-125;

Matches 292; Conservative 38; Mismatches 55; Indels 79; Gaps 6;

QY 1 MARVLGAPVALGLMSLCMSLAIAATPLPTSA--HGVAAGT--TKRDPVTERCSGMSGSDA 58

DB 1 MARVALANTLVGLGLMSLAVALAPLP--TAKGVAAGVETKSDSVPEHCLDTWSPDA 58

QY 59 TLLDNGTMLFPKGEFVWKSHKMDRELISERKNPSPVDAAPROGNSVFLIKGDYKVV 118

DB 59 ATMDHNGTMLFPKGEFVWRGHSCTRELISAKMKPIISVDAAP--GPDVFLIKEDYKVV 117

QY 119 YPPEKKEKGYPKLQDSEPPGIPSPPLDAVBCHEGCEQAEVGLFTPOGH----- 165

DB 118 YPPEKKEKNGYPKLQDSEPPGIPSPPLDAVBCHEGCEQAEVGLFTPOGH----- 165

QY 166 -----GH-- 167

DB 178 BRKSTVGNCTALAKRLERYCFQGNKPLRNPVTGVPVPPYPLDADYVSCPGHGR 237

QY 168 -RNGTGHGNGTTHQPEYWRCSPLVLALTSNNGATYAFSGTHYRMLDTSRDGMSMPI 226

DB 238 PRNGTANGNTH--PMHSRCSPDGLTALSDHGATYAFSGTHYRMLDTSRDGMSMPI 295

QY 227 AHQWPOQPSAVDAAPSWBEKLYVQGTQVYVFLTKGTYTVLSGYPKLEKEVGTPHGII 286

DB 296 AHQWPOQPSAVDAAPSWBEKLYVQGTQVYVFLTKGTYTVLSGYPKLEKEVGTPHGII 286

QY 287 DSVDAATCGSSASRLHIMAGRLWMLDLKGAQATWTELPHPHAKDGLCMESLSPNS 346

DB 356 ETLDAASCRSSSLVYSSGRLWMLDLKGAQATWTELPHPHAKDGLCMESLSPNS 346

QY 347 CSANGPGLYLHGNLYCYSPVEKLANAKALPOPONTSLIGCT 390

DB 416 CSANGSGLYPIHGNLYCYSPVEKLANAKALPOPONTSLIGCT 390

RESULT 5

HEMO_PIG STANDARD, PRT, 459 AA.

AC P50828;

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DB Hemopexin precursor (Hyaluronidase) (BC 3.2.1.35).

HPX.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-48 AND 368-388.

RC TISSUE=Liver;

RA MEDLINE=95096047; Pubmed=7798203;

RA Zhu L., Hope T.J., Hall J., Davies A., Stern M.,

RA Mueller-Berthard U., Stern R., Parslow T.G.;

RT "Molecular cloning of a mammalian hyaluronidase reveals identity with

RT hemopexin, a serum heme-binding protein."

RL J. Biol. Chem. 269:32092-32097(1994).

CC -1- FUNCTION: Binds heme and transports it to the liver for breakdown

CC and iron recovery, after which the free hemopexin returns to the

CC circulation.

CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1->4-linkages between N-

CC acetyl-beta-D-glucosamine and D-glucuronate residues in

CC hyaluronate.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.

CC -1- SIMILARITY: Contains 5 hemopexin-like domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC

DR EMBL, U14751; AAC48457.1; .

DR F1R; A55486; A55486.

DR HSP; P20058; IHN.

DR InterPro; IPR000585; Hemopexin.

DR Pfam; PF00045; hemopexin; 5.

DR SMART; SM00120; HX; 5.

DR PROSITE; PS00024; HEMOPEXIN; 2.

DR Glycoprotein; Heme; Plasma; Repeat; Transport; Signal; Hydrolase;

KW Glycosidase.

FT SIGNAL 1 28

FT CHAIN 29 459 HEMOPEXIN.

FT DOMAIN 58 95 HEMOPEXIN-LIKE 1.

FT DOMAIN 99 142 HEMOPEXIN-LIKE 2.

FT DOMAIN 189 232 HEMOPEXIN-LIKE 3.

FT DOMAIN 256 299 HEMOPEXIN-LIKE 4.

FT DOMAIN 301 344 HEMOPEXIN-LIKE 5.

FT METAL 81 81 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

FT METAL 151 151 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

FT DISULFID 52 232 BY SIMILARITY.

FT DISULFID 150 155 BY SIMILARITY.

FT DISULFID 189 201 BY SIMILARITY.

FT DISULFID 250 453 BY SIMILARITY.

FT DISULFID 359 401 BY SIMILARITY.

FT DISULFID 411 428 BY SIMILARITY.

FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 459 AA; DB06B44C29789CF CRC64;

Query Match 70.7%; Score 1544; DB 1; Length 459;

Best Local Similarity 63.4%; Pred. No. 6.8e-125;

Matches 289; Conservative 33; Mismatches 64; Indels 70; Gaps 5;

QY 1 MARVLGAPVALGLMSLCMSLAIAATPLPTSA--HGVAAGT--TKRDPVTERCSGMSGSDA 58

DB 1 MARVLGAPVALGLMSLCMSLAIAATPLPTSA--HGVAAGT--TKRDPVTERCSGMSGSDA 58

QY 59 TLLDNGTMLFPKGEFVWKSHKMDRELISERKNPSPVDAAPROGNSVFLIKGDYKVV 118

DB 61 TLLDNGTMLFPKGEFVWKSHKMDRELISERKNPSPVDAAPROGNSVFLIKGDYKVV 120

THIS PAGE BLANK (user)

LOCUS	AV653336	631 bp	mRNA	linear	EST 15-JAN-2002
AV653336	GLC Homo sapiens cDNA clone GLCDH02 3', mRNA sequence.				
DEFINITION	AV653336				
ACCESSION	AV653336				
VERSION	AV653336.1	GI:98744350			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 631)				
AUTHORS	Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,H., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.				
JOURNAL	Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver				
MEDLINE	21625106				
PUBMED	11752456				
COMMENT	Contact: Zeguana Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@hgc.sh.cn This clone is available at CHGC in Shanghai.				
FEATURES	Location/Qualifiers				
SOURCE	1..631 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone=GLCDH02 /issue_type="corresponding non cancerous liver tissue" /dev_stage="Adult" /lab_host="SOLR" /clone_lib="GLC" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"				
BASE COUNT	138 a 180 c 161 g 149 t 3 others				
ORIGIN					
Query Match	11.4% Score 425.6; DB 9; Length 631;				
Best Local Similarity	98.4%; Pred. No. 1.1e-95;				
Matches 428; Conservative	7; Indels 0; Gaps 0;				
QY	1479 CAGCAGCGCGGCTGTGTGCTGTGACCTGAGTCTGAGAGCCCAAGCCAGTGTGACAGAGC	1538			
Db	106 CAGCAGCGCGGCTGTGTGCTGTGACCTGAGTCTGAGAGCCCAAGCCAGTGTGACAGAGC	165			
QY	1539 TTCCTTGGCCCCCATGAGAAAGGTAGACGAGGCTTGTGTATGAAAAGTCCCTTGGCCCTTA	1598			
Db	166 TTCCTTGGCCCCCATGAGAAAGGTAGACGAGGCTTGTGTATGAAAAGTCCCTTGGCCCTTA	225			
QY	1599 ACTCATGTTCCGCGCAATGATGTCCTCCGCTTGTAACCTCATCAATGTCCTCCAAATTGTATGCTT	1658			
Db	226 ACTCATGTTCCGCGCAATGATGTCCTCCGCTTGTAACCTCATCAATGTCCTCCAAATTGTATGCTT	285			
QY	1659 ACAGTGTATGTGAGAACTGATGATGAGCCCAAGGCCCTTCCGCAACCCAGATGTGACCA	1718			
Db	286 ACAGTGTATGTGAGAACTGATGATGAGCCCAAGGCCCTTCCGCAACCCAGATGTGACCA	345			
QY	1719 GTCTTCCTGGGCTGCACTCACTGAGGGGCTTCTGACATGATCTTGGCTGGCCCACTTC	1778			
Db	346 GTCTTCCTGGGCTGCACTCACTGAGGGGCTTCTGACATGATCTTGGCTGGCCCACTTC	405			
QY	1779 CTAGTCTCTCAATAAAGACAGATGTGCTCTTGCTCTTCACTGAGGGGCTTCTGACA	1838			
Db	406 CTAGTCTCTCAATAAAGACAGATGTGCTCTTGCTCTTCACTGAGGGGCTTCTGACA	465			

Dy		1899	TTCGATCGGGCTGGAGCCACCTCCCACTTGTGCATATAAAGCAGAATTCTTTCCAC	1899
Dd		466	TGACTTGCGGCTGGCCCCCACCTCCCAGTTTCTCATATNAAACAAGTTCTTTCCAC	525
Dy		1899	TTGTAATCAAAGGACC	1913
Dd		526	TTGAAAANAANNAAAC	540
RESULT 7				
AA705094/c				
LOCUS				
DEFINITION				
AA705094				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
Homo sapiens	(human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
1 (bases 1 to 548)				
Hillier L., Allen M., Bowles L., Dubugue T., Geisels G., Joet S., Kitzman D., Kucba T., Le N., Lennon G., Marra M., Martin J., Moore B., Schellenberg R., Stepeco M., Tan P., Theisinger B., White Y., Wyllie T., Waterston R. and Wilson R.				
WashU-NCI Human EST Project				
Unpublished				
Contact: Wilison RK				
Washington University School of Medicine				
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
Tel.: 314 286 1800				
Fax: 314 286 1810				
Email: est@wustl.wustl.edu				
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information.				
Seq primer: -40mJ3 fwd, RT from Amershams High quality sequence stop: 462.				
Location/Qualifiers				
1..548				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="GBD:3753522"				
/db_xref="taxon:9606"				
/clone="IMAGE:462499"				
/sex="male"				
/dev_stage="20 week-post conception fetus"				
/lab_host="DH10B (ampicillin resistant)"				
/clone_1b="Soares fetal liver spleen INPLS.S1"				
/note="Organ: Liver and Spleen; Vector: pTV73D (pharmacia) with a modified polylinker site 1: Pac I; Site 2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INPLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGCAAGATTAATTAAGAATCTTTTTTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTV73 vector. Library went through one round of normalization. Library constructed by Bento Soares and W.Fatima Bonaldo."				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity	100.0%	Pred. NO. 4.7e+99;	Length 548;	
Matches 423;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
OY		1479	CAGACGCCGCGCTGTGTGCTGGACCTTGAAGTCAGAGCCCAAGCGTAGCAGAGC	1538
Dd		423	CAGACGCCGCGCTGTGTGCTGGACCTTGAAGTCAGAGCCCAAGCGTAGCAGAGC	364
OY		1539	TTCCTTGAGCCCATGAGAGTCAGAGCGAGCTTGTGTATGAAAAAGTCCCTTGAGCCCTA	1598


```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCEGA10"
/trisue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1lb="GLC"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      105 a      124 c      102 g      102 t
ORIGIN
Query Match      11.2%; Score 419.4; DB 9; Length 433;
Best Local Similarity 99.8%; Pred. No. 3.6e-98;
Matches 420; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1494 GGTGGCTGACCTGAGTACAGAGCCCAAGCCAGTGAAGAGCTTCTTGGCCCATG 1553
DB 1 GGTGGCTGACCTGAGTACAGAGCCCAAGCCAGTGAAGAGCTTCTTGGCCCATG 60

QY 1554 AGAAGGTAGACGAGCCTTGTATGAGAAAGTCCCTTGGCCCTTACTCATGTTCCGCCA 1613
DB 61 AGAAGGTAGACGAGCCTTGTATGAGAAAGTCCCTTGGCCCTTACTCATGTTCCGCCA 120

QY 1614 ATGGTCCCGCTGTACTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673
DB 121 ATGGTCCCGCTGTACTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180

QY 1674 AACTGAATGACGAGCCCTTCCGCAACCCAGAAATGAGACCAAGTCTCTGGGCTGCA 1733
DB 181 AACTGAATGACGAGCCCTTCCGCAACCCAGAAATGAGACCAAGTCTCTGGGCTGCA 240

QY 1734 CTCACTAGAGGGGCTTCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1793
DB 241 CTCACTAGAGGGGCTTCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 300

QY 1794 AAGACAGATTGCTTCTGCTCTCACTGAGGGGCTTCTGACATGATGATGATGATGATGAT 1853
DB 301 AAGACAGATTGCTTCTGCTCTCACTGAGGGGCTTCTGACATGATGATGATGATGATGAT 360

QY 1854 CCCACCTCCCAAGTTCTCATATTAAGACAGATTGCTTCTTCACTTGAATCAAGGAGCC 1913
DB 361 CCCACCTCCCAAGTTCTCATATTAAGACAGATTGCTTCTTCACTTGAATCAAGGAGCA 420

QY 1914 T 1914
DB 421 T 421

RESULT 10
AV655401 433 bp mRNA linear EST 15-JAN-2002
LOCUS AV655401 GLC Homo sapiens cDNA clone GLCEGC04 3', mRNA sequence.
DEFINITION AV655401
ACCESSION AV655401
VERSION AV655401.1 GI:9876415
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 433)
XU,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z., and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL MEDLINE 11752456
PUBMED 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
```

```
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. 433
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCEGC04"
/trisue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1lb="GLC"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      106 a      124 c      102 g      101 t
ORIGIN
Query Match      11.2%; Score 419; DB 9; Length 433;
Best Local Similarity 100.0%; Pred. No. 4.6e-98;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1494 GGTGGCTGACCTGAGTACAGAGCCCAAGCCAGTGAAGAGCTTCTTGGCCCATG 1553
DB 1 GGTGGCTGACCTGAGTACAGAGCCCAAGCCAGTGAAGAGCTTCTTGGCCCATG 60

QY 1554 AGAAGGTAGACGAGCCTTGTATGAGAAAGTCCCTTGGCCCTTACTCATGTTCCGCCA 1613
DB 61 AGAAGGTAGACGAGCCTTGTATGAGAAAGTCCCTTGGCCCTTACTCATGTTCCGCCA 120

QY 1614 ATGGTCCCGCTGTACTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673
DB 121 ATGGTCCCGCTGTACTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180

QY 1674 AACTGAATGACGAGCCCTTCCGCAACCCAGAAATGAGACCAAGTCTCTGGGCTGCA 1733
DB 181 AACTGAATGACGAGCCCTTCCGCAACCCAGAAATGAGACCAAGTCTCTGGGCTGCA 240

QY 1734 CTCACTAGAGGGGCTTCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1793
DB 241 CTCACTAGAGGGGCTTCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 300

QY 1794 AAGACAGATTGCTTCTGCTCTCACTGAGGGGCTTCTGACATGATGATGATGATGATGAT 1853
DB 301 AAGACAGATTGCTTCTGCTCTCACTGAGGGGCTTCTGACATGATGATGATGATGATGAT 360

QY 1854 CCCACCTCCCAAGTTCTCATATTAAGACAGATTGCTTCTTCACTTGAATCAAGGAGCC 1912
DB 361 CCCACCTCCCAAGTTCTCATATTAAGACAGATTGCTTCTTCACTTGAATCAAGGAGC 419

RESULT 11
AL534854/c 1201 bp mRNA linear EST 12-MAY-2003
LOCUS AL534854 Homo sapiens PBTAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF006Y122 3'-PRIME, mRNA sequence.
ACCESSION AL534854
VERSION AL534854.2 GI:30541057
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J., and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12798347.
JOURNAL MEDLINE 11752456
PUBMED 11752456
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
```


double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pUT3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 99 a 107 c 123 g 89 t 1 others

Query Match 11.0%; Score 412; DB 14; Length 419;
Best Local Similarity 99.8%; Pred. No. 3.1e-96;
Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1479 CAGACGCGCGCTGTGTGTGCTGAGCTGAGTCAAGTCAAGAGCCCAAGCCAGTGAAGAGC 1538
413 CAGACGCGCGCTGTGTGTGCTGAGCTGAGTCAAGTCAAGAGCCCAAGCCAGTGAAGAGC 354
1539 TTCCTTGGCCCAATGAGAGTGAAGAGCCCTTTGTATGAAAAGTCCCTTGGCCCTTA 1598
353 TTCCTTGGCCCAATGAGAGTGAAGAGCCCTTTGTATGAAAAGTCCCTTGGCCCTTA 294
1599 ACTCATGTTCCGCGCATGATGCTCCGCGCTTGTACCTCATCCATGCTCCCAATTTGTACTGCT 1658
293 ACTCATGTTCCGCGCATGATGCTCCGCGCTTGTACCTCATCCATGCTCCCAATTTGTACTGCT 234
1659 ACAGATGATGAGAACTGATGATGAGCAAGGCGCTTCCGCAACCCAGATGATGACCA 1718
233 ACAGATGATGAGAACTGATGATGAGCAAGGCGCTTCCGCAACCCAGATGATGACCA 174
1719 GTCTCTGGGCTGCTCACTGATGAGGCGCTTGTGACATGATGCTGCGCTTGGCCCACTTC 1778
173 GTCTCTGGGCTGCTCACTGATGAGGCGCTTGTGACATGATGCTGCGCTTGGCCCACTTC 114
1779 CTAGTCTCTCATATATTAAGACAGATGCTTCTTCTGCTTCTCACTGAGGCGCTTGTGACA 1838
113 CTAGTCTCTCATATATTAAGACAGATGCTTCTTCTGCTTCTCACTGAGGCGCTTGTGACA 54
1839 TGAGTCTGGGCTGAGCCCACTCCCGAGTTTCTCATATTAAGACAGATGCT 1891
53 TGAGTCTGGGCTGAGCCCACTCCCGAGTTTCTCATATTAAGACAGATGCT 1

RESULT 13 1053 bp mRNA linear EST 31-MAY-2003
LOCUS AL564329/c
DEFINITION AL564329 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM004YJ20 3-PRIME, mRNA sequence.

ACCESSION AL564329
VERSION AL564329.2 GI:31288307
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1. W.B., Gruber, C., Jessup, J. and Polajnar, D.
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1053)
JOURNAL Full-length cDNA libraries and normalization
COMMENT On Feb 16, 2001 this sequence version replaced gi:12914625.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5958.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM004DB10NP1&cluster=5958.r. Contact :
Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/>
Faraday Avenue Genoscope sequence ID: CS0DM004DB10NP1.

FEATURES
source 1.1053
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004YJ20"
/issue="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-spligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

BASE COUNT 237 a 286 c 306 g 210 t 14 others

Query Match 11.0%; Score 411.6; DB 9; Length 1053;
Best Local Similarity 72.3%; Pred. No. 6.3e-96;
Matches 698; Conservative 2; Mismatches 11; Indels 254; Gaps 3;

928 TCACAGAGGACCACTACTGCGCTGTGACACACGCGGGATGGCTGATAGTGGCC 987
719 TTGATGAGGAGCCCACTGCGCTGTGACACACGCGGGATGGCTGATAGTGGCC 660
988 ATTGCTCATGATGCGCCCAAGGCTCTTCAAGATGATGCTGCTTTCTTGGAGAA 1047
659 ATTGCTCATGATGCGCCCAAGGCTCTTCAAGATGATGCTGCTTTCTTGGAGAA 600
1048 AAATCTATCTGTGCTCAAGTGTGTATTTGGGAGAGGCTTGAAGTGAAGTGGACAG 1107
599 AAATCTATCTGTGCTCAAGTGTGTATTTGGGAGAGGCTTGAAGTGAAGTGGACAG 585
1108 CATATCAACTGTGTATTTATTAACATCTTGTCTCCAGGCGACCCAGATATGCT 1167
584 -----CAGGACCCAGATATATGCT 563
1168 TCCTGACAAAGGAGGCTATACCTAGTAAAGCGTTATCCGAAGCGCTGAGAGGAG 1227
562 TCCTGACAAAGGAGGCTATACCTAGTAAAGCGTTATCCGAAGCGCTGAGAGGAG 503
1228 TCGGAGCCCTCATGAGGATTTATCTGAGCTGTGTGATGCGGCTTTATCTGCGTGGGT 1287
502 TCGGAGCCCTCATGAGGATTTATCTGAGCTGTGTGATGCGGCTTTATCTGCGTGGGT 443
1288 CTTCGCGCTCCATATCATGAGGAGGAGGCTTGTGGGTGCTTGAAGGAGCATTGTT 1347
442 CTTCGCGCTCCATATCATGAGGAGGAGGCTTGTGGGTGCTTGAAGGAGCATTGTT 422
1348 CTGCTACTGTGTGTGATAGTATCCCAACAGAGGATGAGAGGCTTAGTCAAGATCC 1407
421 ----- 422
1408 CCAGGCGATGAGAGGCTTAGTCAAGATCCCAAGATGAGAGGCTTAGTCAAGATCC 1467
421 ----- 422
1468 TGCCTTCTCCAGAGCGCGGCTGTGTGCTGAGCTGAGTCAAGAGCCCAAGCCAC 1527
421 -----CAGGACGCGCGGCTGTGTGCTGAGCTGAGTCAAGAGCCCAAGCCAC 373
1528 GTGACAGAGCTTCTTGGCCCATGAGAGGATGAGAGGCTTGTGATGAGAAAGTC 1587
372 GTGACAGAGCTTCTTGGCCCATGAGAGGATGAGAGGCTTGTGATGAGAAAGTC 313
1588 CTTGGCCCTTAATCATGATGTTCCGCAATGATGCTCCGCTTGTATCATCATGATGCTCCAA 1647
312 CTTGGCCCTTAATCATGATGTTCCGCAATGATGCTCCGCTTGTATCATCATGATGCTCCAA 253
1648 TTGTATCTGCTCAAGATGATGAGAGAACTGATGAGAGGAGGCTTCCGCAAGCCCA 1707
252 TTGTATCTGCTCAAGATGATGAGAGAACTGATGAGAGGAGGCTTCCGCAAGCCCA 193
1708 GAATGACCAAGTCTTCTGAGGCTGACATCACTGAGGAGGCTTCTGACATGATGCTGGCCT 1767
192 GAATGACCAAGTCTTCTGAGGCTGACATCACTGAGGAGGCTTCTGACATGATGCTGGCCT 133

QY 1768 GAGCCCACTCTAGTTCCTCATATAAAGACAGATTGCTTCTCGCTTCACTGAGGG 1827
 DB 132 GAGCCCACTCTAGTTCCTCATATAAAGACAGATTGCTTCTCGCTTCACTGAGGG 73
 QY 1828 GCTTCTGACATGATGCTGCGCTGCGCCACCTCC-CGAGTTTCTCATATAAAGACAGA 1886
 DB 72 GCTTCTGACATGATGCTGCGCTGCGCCACCTCC-CGAGTTTCTCATATAAAGNNNA 13
 QY 1887 TTGCT 1891
 DB 12 WNWMT 8
 RESULT 14
 AA514026 485 bp mRNA linear EST 13-NOV-1997
 LOCUS n132e01.s1 NCI_CGAP_Ew1 Homo sapiens cDNA clone IMAGE:972024
 DEFINITION similar to gb:J03048 HEMOPEXIN PRECURSOR (HUMAN) ; mRNA sequence.
 ACCESSION AA514026
 VERSION AA514026.1 GI:2252447
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-@email.nih.gov
 Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D. Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LML at:
 www.bio.lml.gov/bdtp/image/image.html
 Insert Length: 602 Std Error: 0.00
 Seq primer: -40ml3 fwd. RT from Amersham
 High quality sequence stop: 388.
 FEATURES
 source
 1..485
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:972024"
 /tissue_type="Ewing's sarcoma"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Ew1"
 /note="Vector: PAMPI0; mRNA made from Ewing's sarcoma,
 cDNA made by oligo-dT priming. Non-directionally cloned.
 Size-selected on agarose gel, average insert size 600 bp.
 Reference: Krizman et al. (1996) Cancer Research
 56:5380-5383."
 BASE COUNT 96 a 143 c 124 g 122 t
 ORIGIN
 Query Match 11.0%; Score 411; DB 9; Length 485;
 Best Local Similarity 99.8%; Pred. No. 6,1e-96;
 Matches 422; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1479 CAGAGCGCGCTGTGTGCTGAGCTGAGCTGAGCCCAAGCCAGCTGAGCAGAGC 1538
 DB 64 CAGAGCGCGCTGTGTGCTGAGCTGAGCTGAGCCCAAGCCAGCTGAGCAGAGC 123
 QY 1539 TTCTCTGCGCCCATGAGAGTAGACGAGGCTTGTATGAAAAGTCCCTTGCCCTTA 1598
 DB 124 TTCTCTGCGCCCATGAGAGTAGACGAGGCTTGTATGAAAAGTCCCTTGCCCTTA 183

QY 1599 ACTCATGTTCCGCAATGATGTCGCCGCTGTACTCATCCATGATGTCCTTACTGCT 1658
 DB 184 ACTCATGTTCCGCAATGATGTCGCCGCTGTACTCATCCATGATGTCCTTACTGCT 243
 QY 1659 ACAATGATGTGGAGAAATGTAATGCAAGGCCCTTCCGCAACCCAGAAATGTGACCA 1718
 DB 244 ACAATGATGTGGAGAAATGTAATGCAAGGCCCTTCCGCAACCCAGAAATGTGACCA 303
 QY 1719 GTCTCCGCGGCTGCACTCACTGAGGGGCTTGTGACATGATGCTGAGGCTGCCCCACCTC 1778
 DB 304 GTCTCCGCGGCTGCACTCACTGAGGGGCTTGTGACATGATGCTGAGGCTGCCCCACCTC 362
 QY 1779 CTAGTTCCTCATATAAAGACAGATTGCTTCTCGCTTCACTGAGGGGCTTGTGACA 1838
 DB 363 CTAGTTCCTCATATAAAGACAGATTGCTTCTCGCTTCACTGAGGGGCTTGTGACA 422
 QY 1839 TGAATCTGCGCTTGGCCCACTCCCAAGTTTCTCATATAAAGACAGATTGCTTCTCAC 1898
 DB 423 TGAATCTGCGCTTGGCCCACTCCCAAGTTTCTCATATAAAGACAGATTGCTTCTCAC 482
 QY 1899 TTG 1901
 DB 483 TTG 485
 RESULT 15
 A1283497/c 494 bp mRNA linear EST 27-JAN-1999
 LOCUS qn194e01.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:1854648 3' similar to gb:J03048 HEMOPEXIN PRECURSOR (HUMAN)
 ; contains element TARI repetitive element ; mRNA sequence.
 ACCESSION A1283497
 VERSION A1283497.1 GI:3921730
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-@email.nih.gov
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.lml.gov) for further information.
 Insert Length: 3834 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 380.
 FEATURES
 source
 1..494
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1854648"
 /lab_host="DH10B"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /note="Organ: pooled; Vector: pT7TD-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19, testis NHT, and B-cell
 NCI CGAP GCBI) were mixed and as circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The diver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 119 a 126 c 145 g 104 t
 ORIGIN

Query Match	10.9%	Score 406.6;	DB 9;	Length 494;
Best Local Similarity	99.0%;	Pred. No. 8.7e-35;		
Matches 409; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1479	CAGGACGGGGCTGTGTGGCTGTGACCTGGAAGTCAAGAGGCCCAAGCACAAGTGAACAAGC	1538
Db	413	CAGGACGGGGCTGTGTGGCTGTGACCTTAAAGTCAAGAGGCCCAAGCACAAGTGAACAAGC	354
QY	1539	TTCTTGGCCCATGAGAAAGGTAGAAGGAGCCTTGTGTATGAAAAAGTCCCTTGGCCCTA	1598
Db	353	TTCTTGGCCCATGAGAAAGGTAGAAGGAGCCTTGTGTATGAAAAAGTCCCTTGGCCCTA	294
QY	1599	ACTATATGTTCCGCCAATATGTCTCCGGCTTGTACCTTCATCCATGATGTCCTCAATTGTATGCT	1658
Db	293	ACTATATGTTCCGCCAATATGTCTCCGGCTTGTACCTTCATCCATGATGTCCTCAATTGTATGCT	234
QY	1659	ACAGTGAATGTGAGAAACTGAATGAGCAAGAGGCCCTTCCGCAACCCCAAGAAATGTAGCA	1718
Db	233	CAAGTGAATGTGAGAAACTGAATGAGCAAGAGGCCCTTCCGCAACCCCAAGAAATGTAGCA	174
QY	1719	GTCCTCTGGGCTGACATCACTGAGGGGCTTCTGACATGATGTGGCTGTGCCCCCACTC	1778
Db	173	GTCCTCTGGGCTGACATCACTGAGGGGCTTCTGACATGATGTGGCTGTGCCCCCACTC	114
QY	1779	CTAGTCTCTAATAATAAGACAGATGTCTTCTTGCTTCTCACTGAGGGGCTTCTTGACA	1838
Db	113	CTAGTCTCTAATAATAAGACAGATGTCTTCTTGCTTCTCACTGAGGGGCTTCTTGACA	54
QY	1839	TGAGTCTGGGCTGGGCCCACTCCCAAGTTCCTCATATAATAAGACAGATGTCT	1891
Db	53	TGAGTCTGGGCTGGGCCCTCCCAAGTTCCTCATATAATAAGACAGATGTCT	1

Search completed: December 16, 2003, 05:34:05
Job time : 4682.61 secs

THIS PAGE BLANK (USPTO)

QY 1233 ACCCTCATGGATATCTGACTGCTGTGGATGCGGCTTTATCTGCGGCTTCT 1292
 DB 518 ACCCTCATGGATATCTGACTGCTGTGGATGCGGCTTTATCTGCGGCTTCT 577
 QY 1293 CGGCTCATATCATGCGAGGTGAAGGCT 1321
 DB 578 CGGCTCATATCATGCGAGGTGAAGGCT 606
 RESULT 3
 BX325177/c 992 bp mRNA linear EST 02-MAY-2003
 LOCUS BX325177 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0D010YE24 3-PRIME, mRNA sequence.
 ACCESSION BX325177
 VERSION BX325177.1 GI:30340447
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 992)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqef@genoscope.cns.fr, web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5958.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0A010BC12NP1&cluster=5958.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0A010BC12NP1.
 Location/Qualifiers
 1. 992
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D010YE24"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 228 a 270 c 293 g 199 t 2 others
 ORIGIN
 Query Match 11.5%; Score 429.8; DB 13; Length 992;
 Best Local Similarity 73.3%; Pred. No. 1.1e-100; Indels 253; Gaps 2;
 Matches 704; Conservative 1; Mismatches 3;
 QY 928 TCCACAGGAGCCACTACTGCTGCTGACACACGCGGATGCTGACATAGTGGCC 987
 DB 708 TTCAGTGGAGCCACTACTGCTGCTGACACACGCGGATGCTGACATAGTGGCC 649
 QY 988 ATTGCTCATCATGTCGCCCGGAGGCTCTTCAGCATGAGTGTGCTTTCTGGAGAA 1047
 DB 648 ATTGCTCATCATGTCGCCCGGAGGCTCTTCAGCATGAGTGTGCTTTCTGGAGAA 589
 QY 1048 AAACCTATCTGTCCAGGTGTGTATTGGGGAGAGGCTTGAGTAGACACTGGACAAG 1107
 DB 588 AAACCTATCTGTG----- 576
 QY 1108 CATATCAACTCTGTATTATATACATCTTTGCTCCAGGAGCCAGGTATATGTCT 1167

DB 575 -----TCCAGGAGCCACAGTATATGTCT 552
 QY 1168 TCCTGACAAAGGAGGCTTATACCTAGTAAAGGCTTATCCGAAGCGGCTGGAAAGAG 1227
 DB 551 TCCTGACAAAGGAGGCTTATACCTAGTAAAGGCTTATCCGAAGCGGCTGGAAAGAG 492
 QY 1228 TCGGAGACCCCTCATGAGGATTTATCTGAGTCTGTGATGCGGCTTTATCTGCGCTGGGT 1287
 DB 491 TCGGAGACCCCTCATGAGGATTTATCTGAGTCTGTGATGCGGCTTTATCTGCGCTGGGT 432
 QY 1288 CTTCGCGCTCCATATCATATGAGGAGGCTTCTGAGGCTTCTGAGGAGCCTTGT 1347
 DB 431 CTTCGCGCTCCATATCATATG----- 411
 QY 1348 CTGCTACTGTCTGTGGCATATATCCCAACAGGAGCATGAAAGGCTAGTACAGATCC 1407
 DB 410 ----- 411
 QY 1408 CCAGGAGATGAGAGGCTTAGTCCAGATCCCATGACATGAAAGCATGTATTGG 1467
 DB 410 ----- 411
 QY 1468 TGCCTTCTCCAGAGAGCGGCTGTGTGCTGACCTGAACTGAGAGCCCAAGCCAC 1527
 DB 410 -----CAGACGCGGCTGTGTGCTGACCTGAACTGAGAGCCCAAGCCAC 362
 QY 1528 GTGAGACAGGCTTCTTGGCCCCCATGAGAGGATGACGAGGCTTGTATGAAAGATC 1587
 DB 361 GTGAGACAGGCTTCTTGGCCCCCATGAGAGGATGACGAGGCTTGTATGAAAGATC 302
 QY 1588 CTTTGGCCCTTAATCTCATGTTCGCGCAATGATGCCGCTTGTACTCATATGATGCTCCAA 1647
 DB 301 CTTTGGCCCTTAATCTCATGTTCGCGCAATGATGCCGCTTGTACTCATATGATGCTCCAA 242
 QY 1648 TTTGATCTGCTACATGATATGAGAGAACTGAAATGAGGCAAGGCTTCCGCAACCCA 1707
 DB 241 TTTGATCTGCTACATGATATGAGAGAACTGAAATGAGGCAAGGCTTCCGCAACCCA 182
 QY 1708 GAATGTGACAGTCTCTGCGGCTGACCTGACCTGAGAGGCTTCTGACATGATGCTGGCT 1767
 DB 181 GAATGTGACAGTCTCTGCGGCTGACCTGACCTGAGAGGCTTCTGACATGATGCTGGCT 122
 QY 1768 GGCCCCACCTCTAGTCTCTCAATTAATGAAGAGATGCTTCTGCTCTGACATGAGG 1827
 DB 121 GGCCCCACCTCTAGTCTCTCAATTAATGAAGAGATGCTTCTGCTCTGACATGAGG 62
 QY 1828 GCTTCTGACATGATGCTGCGGCTGAGCCCACTCCCACTTCTCATTAATGAAGAGAT 1887
 DB 61 GCTTCTGACATGATGCTGCGGCTGAGCCCACTCCCACTTCTCATTAATGAAGAGAT 2
 QY 1888 T 1888
 DB 1 T 1
 RESULT 4
 CB154152
 LOCUS CB154152 551 bp mRNA linear EST 29-JAN-2003
 DEFINITION K-EST0211925 L17N670205 Homo sapiens cDNA clone L17N670205-4-C01
 5', mRNA sequence.
 ACCESSION CB154152
 VERSION CB154152.1 GI:28139149
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 551)
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 21C Frontier Korean EST Project 2001
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished

Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

Source

1. 632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCDM909"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1lb="GIC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 225 a 97 c 164 g 145 t 1 others
ORIGIN

Query Match 15.6%; Score 583.4; DB 9; Length 632;
Best Local Similarity 98.7%; Pred. No. 5.7e-141;
Matches 598; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

34 CAGATTGGAAGTCATCAGCATATATAGATGCAATTTAAAGCCCTGAGCTAGAGTATAC 93
1 CAGATTGGAAGTCATCAGCATATATAGATGCAATTTAAAGCCCTGAGCTAGAGTATAC 60
94 CAGGAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 153
61 CAGGAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
154 GATATCAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 213
121 GATATCAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
214 TGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 273
181 TGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
274 CAGTGTCTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 333
241 CAGTGTCTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
334 TTGAGAGATCTAGCAACTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
301 TTGAGAGATCTAGCAACTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
394 AATTGAAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 453
361 AATTGAAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
454 TTGTGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 513
421 TTGTGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
514 ACCTCATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
481 ACCTCATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
574 TTCTTTTAAAGTTCTGAGACCTGATATTCATATCTAATACCGCATCTCTGTACATGA 633
540 TTCTTTTAAAGTTCTGAGACCTGATATTCATATCTAATACCGCATCTCTGTACATGA 599
634 CACCTG 639
600 CACCTG 605

RESULT 2
LOCUS AM965111 672 bp mRNA linear EST 01-JUN-2000
DEFINITION EST377184 MAGE resequences, MAGI Homo sapiens cDNA, mRNA sequence.
ACCESSION AM965111
VERSION AM965111.1 GI:8154947

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaepard, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and Quackenbush, J.

TITLE
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray

JOURNAL
COMMENT
Unpublished
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johng@tigr.org
Plate: 210

FEATURES

Source

1. 672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_1lb="MAGE resequences, MAGI"
/note="Vector: pBluescriptSKm"

BASE COUNT 128 a 205 c 167 g 172 t
ORIGIN

Query Match 13.6%; Score 506.6; DB 10; Length 672;
Best Local Similarity 87.4%; Pred. No. 7.2e-121;
Matches 602; Conservative 0; Mismatches 4; Indels 83; Gaps 1;

633 ACACTGAATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
1 ACACTGAATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
693 TAGCCATGTTTTCACACTTGTCTTTTCAAGGACCTGCTGATTCAGTTCAAGCA 752
61 TAGCCATGTTTTCACACTTGTCTTTTCAAGGACCTGCTGATTCAGTTCAAGCA 120
753 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
121 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
813 CTTCATCTTGGGCTCTGAGATGATCTCTCCCTCATAGACTCATTTTGTGATGATCA 872
181 CTTCATCTTGGGCTCTGAGATGATCTCTCCCTCATAGACTCATTTTGTGATGATCA 240
873 CTAGCCTCTTCTGCTGCTGGGCTTCTGCGAGCGGCTTGAAGCACTTATGATTCAC 932
241 CTAGCCTCTTCTGCTGCTGGGCTTCTGCGAGCGGCTTGAAGCACTTATGATTCAC 300
933 AGGAGCCCATCTATGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 992
301 AGGAGCCCATCTATGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
993 TCATCAGTGGGCCCCAGAGGCTCTTCAAGCAGTGAATGCTCTTTCTGGAGAGAAAACT 1052
361 TCATCAGTGGGCCCCAGAGGCTCTTCAAGCAGTGAATGCTCTTTCTGGAGAGAAAACT 420
1053 CTATCTGCTTCAGAGTGTATTTGGGGAGAGGCTTGAAGTAGAGACTGGAGCAAGATAT 1112
421 CTATCTGG----- 428
1113 CCAACTGTGATTTATTAACATCTTGTCTCCAGAGGACCCAGGATATATGCTTCCTG 1172
429 -----TCCAGGAGACCCAGGATATATGCTTCCTG 457
1173 ACAAGAGAGAGGCTATACCTTAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
458 ACAAGAGAGAGGCTATACCTTAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 10:53:38 / Search time 4675.94 seconds
(without alignments)
19424.068 Million cell updates/sec

Title: US-09-900-448-3_COPY_10001_13737

Perfect score: 3737
Sequence: 1 cccaaggagaggtcagtcac.....acctttgtgtgttcac 3737

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estc:*
10: gb_estc2:*
11: gb_estc3:*
12: gb_estc4:*
13: gb_estc5:*
14: gb_estc6:*
15: em_estfun:*
16: em_estcom:*
17: em_esthum:*
18: em_estinv:*
19: em_estpln:*
20: em_estvrt:*
21: em_estfun:*
22: em_estmam:*
23: em_estmus:*
24: em_estpro:*
25: em_estrod:*
26: em_estphg:*
27: em_estvrt:*
28: gb_estc1:*
29: gb_estc2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	583.4	15.6	632	9	AV653618
2	506.6	13.6	672	10	AW965111 EST377184
3	429.8	11.5	992	13	BX325177 BX325177
4	425.8	11.4	551	14	CB154152 CB154152 K-EST0211

c	5	425.8	11.4	704	9	AV645403	AV645403
c	6	425.6	11.3	631	9	AV653336	AV653336
c	7	423	11.3	548	9	AA705094	AA705094 2193110.8
c	8	419.8	11.2	562	9	AT798878	AT798878 w93508.x
c	9	419.4	11.2	433	9	AV655383	AV655383
c	10	419	11.2	433	9	AV655401	AV655401
c	11	416.2	11.1	1201	9	AL534854	AL534854
c	12	412	11.0	419	14	RI7061	RI7061 yf45a11.e2
c	13	411.6	11.0	1053	9	AL564329	AL564329
c	14	411	11.0	485	9	AA514026	AA514026 n132e01.8
c	15	406.6	10.9	494	9	AT283497	AT283497 qh94e01.x
c	16	406.4	10.9	465	14	RS3514	RS3514 yf70f12.87
c	17	402.6	10.8	914	13	BX412777	BX412777 BX412777
c	18	401	10.7	401	9	AI204645	AI204645 qd44a06.x
c	19	397.8	10.6	1201	9	AL564746	AL564746
c	20	396	10.6	460	14	T74878	T74878 yc58b10.81
c	21	394	10.5	545	14	T54902	T54902 yd45b02.81
c	22	386	10.3	462	14	HS1314	HS1314 yd49g03.81
c	23	386	10.3	503	14	T71372	T71372 yc61h04.81
c	24	384.4	10.3	509	14	T64589	T64589 yc25b12.81
c	25	380.2	10.2	1201	13	BX443958	BX443958
c	26	378	10.1	877	10	BG196922	BG196922 RST16153
c	27	375.8	10.1	483	14	T62823	T62823 yc70b10.81
c	28	371.8	9.9	446	14	T53805	T53805 yd85a03.81
c	29	371.6	9.9	429	9	AV719518	AV719518
c	30	370	9.9	442	14	H69341	H69341 yu19c10.81
c	31	369.8	9.9	476	9	AV718855	AV718855
c	32	365	9.8	416	14	T68282	T68282 yc41b03.81
c	33	365	9.8	433	14	T69134	T69134 yc32c08.81
c	34	364.8	9.8	1201	13	BX405876	BX405876
c	35	363	9.7	474	14	T71683	T71683 yc62h05.81
c	36	362.4	9.7	440	14	H62225	H62225 yf15e01.81
c	37	360.4	9.6	512	14	T72898	T72898 yc51b05.81
c	38	360	9.6	512	14	T67915	T67915 yc28b02.81
c	39	352.8	9.4	408	14	T74698	T74698 yc6f03.81
c	40	352.6	9.4	472	14	NS2420	NS2420 yv50f12.81
c	41	351.8	9.4	1201	13	BX440692	BX440692
c	42	350.8	9.4	359	14	H81206	H81206 yu99e11.81
c	43	344.8	9.2	393	14	T62830	T62830 yc70c06.81
c	44	336.4	9.0	432	9	AI076380	AI076380 c209a06.x
c	45	333	8.9	488	14	T69464	T69464 yc38h04.81

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
AV653618	AV653618	GLC Homo sapiens cDNA clone GLCDMG09 3', mRNA sequence.	AV653618	AV653618.1	GI:9874632	EST.	Homo sapiens (human)	1 (bases 1 to 632)	Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.	Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver	Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)	11752456	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45)

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(488)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30754

Query Match 4.0%; Score 149.4; DB 11; Length 488;
Best Local Similarity 73.4%; Pred. No. 2.3e-37;
Matches 246; Conservative 0; Mismatches 6; Indels 83; Gaps 1;

QY 928 TCACAGGAGGACCACTACTGCGCTGAGACACCGCGGATGCTGGATAGTGGCCC 987
DB 237 TTCAGTGGAGGACCACTACTGCGCTGAGACACCGCGGATGCTGGATAGTGGCCC 236
QY 988 ATTGCTCATCAGTGGCCCGGAGGCTCTTCAGCAGTGAATGCTCTTTCTGGAGAA 1047
DB 297 ATTGCTCATCAGTGGCCCGGAGGCTCTTCAGCAGTGAATGCTCTTTCTGGAGAA 356
QY 1048 AAATCTATCTGGTCCAGGTGTGTATATGGGAGAGGCTTGAAGTATAGACTGGACAA 1107
DB 357 AAATCTATCTGGTCCAGGTGTGTATATGGGAGAGGCTTGAAGTATAGACTGGACAA 371
QY 1108 CATATCCAACTCTGTATTATTATTAACATCCTTTGCTCCAGGACCCAGATATATGCT 1167
DB 372 -----CAGGGACCCAGATATATGCT 393
QY 1168 TCCTGCAAGAGGAGGCTATACCTTAGTAGCGGTTATCCGAAGCGGCTGGAAGAG 1227
DB 394 TCCTGCAAGAGGAGGCTATACCTTAGTAGCGGTTATCCGAAGCGGCTGGAAGAG 453
QY 1228 TCGGAGCCCTCATGGGATATATCTGACTCTGTG 1262
DB 454 ACGGAGCCCTCATGGGATATATCTGACTCTGTG 488

RESULT 12
US-09-918-995-32181
Sequence 32181, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 32181
LENGTH: 473
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(473)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32181

Query Match 3.6%; Score 134.2; DB 11; Length 473;
Best Local Similarity 97.8%; Pred. No. 2.1e-32;
Matches 136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 928 TCACAGGAGGACCACTACTGCGCTGAGACACCGCGGATGCTGGATAGTGGCCC 987
DB 323 TTCAGTGGAGGACCACTACTGCGCTGAGACACCGCGGATGCTGGATAGTGGCCC 382
QY 988 ATTGCTCATCAGTGGCCCGGAGGCTCTTCAGCAGTGAATGCTCTTTCTGGAGAA 1047
DB 383 ATTGCTCATCAGTGGCCCGGAGGCTCTTCAGCAGTGAATGCTCTTTCTGGAGAA 442

QY 1048 AAATCTATCTGGTCCAGG 1066
DB 443 AAATCTATCTGGTCCAGG 461

RESULT 13
US-09-918-995-32820
Sequence 32820, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 32820
LENGTH: 491
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(491)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32820

Query Match 3.6%; Score 134.2; DB 11; Length 491;
Best Local Similarity 97.8%; Pred. No. 2.2e-32;
Matches 136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 928 TCACAGGAGGACCACTACTGCGCTGAGACACCGCGGATGCTGGATAGTGGCCC 987
DB 272 TTCAGTGGAGGACCACTACTGCGCTGAGACACCGCGGATGCTGGATAGTGGCCC 331
QY 988 ATTGCTCATCAGTGGCCCGGAGGCTCTTCAGCAGTGAATGCTCTTTCTGGAGAA 1047
DB 332 ATTGCTCATCAGTGGCCCGGAGGCTCTTCAGCAGTGAATGCTCTTTCTGGAGAA 391
QY 1048 AAATCTATCTGGTCCAGG 1066
DB 392 AAATCTATCTGGTCCAGG 410

RESULT 14
US-09-960-352-10323
Sequence 10323, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathalegan, Nagappan
TITLE OF INVENTION: NOVEL NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
PRIOR FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 10323
LENGTH: 420
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 44-LIB34-041-Q1-B1-C8
US-09-960-352-10323

Query Match 3.3%; Score 123.6; DB 10; Length 420;
Best Local Similarity 79.0%; Pred. No. 5.8e-29;
Matches 147; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 28-LIB34-031-Q1-E1-G7
US-09-960-352-6580

Query Match 5.3%; Score 197.4; DB 10; Length 400;
Best Local Similarity 75.8%; Pred. No. 3.9e-53;
Matches 257; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

1479 CAGAGCGCGGCTGTGTGCTGCTGACCTTGAAGTCAGAGAGCCCAAGCCAGCTGACAGAGC 1538
111 CAGGCCAGAGCTGTCAAGCTGTGACCTTGAAGTCTGAAGAGCTCAAGCCAGCTGACAGAGA 307
1539 TTCCTTGGCCCGCATGAGAGGTAGAGAGGCTTGTATGGAAGATCCCTTGGCCCTTA 1598
306 TTCCTTGGCTCCATACGAAAGTGTGATGGGCTCTGTATGAGAGATATCTGGCCCCC 247
1599 ACTCATGTTCGGCAATGTCCCGCTTGTACTCATCTCATGTGTCCCAATTTGTACTGCT 1658
246 ACTCGTATGTGCAATGTCTGGGCTTGTACCTCGTCAAGGCCCAATCTGTACTGCT 187
1659 ACAGTATGTGAGAACTGATGACAGCCAGGCCCTTCCGCAACCCCAAGATGTGACCA 1718
186 ACAAGATGTGAGAGATTGAGACGACCAAGGACTTCCGCAAGGCCCAAGAGATGAACA 127
1719 GTCTCCGGGCTGCACTGACCTGAGAGGCGCTTTCGACATGATGCTGGCCCACTC 1778
126 GCCTCTGGGCTGCTGCTCCCAAGCAGCACTCTGATGCGAGTTTGGCTCAGATCTCTTC 67
1779 CTAG-TTCCTCATATTAAGACAGATTGCTTCTGCTT 1816
66 CCAATTTCATCATATTAAGCAGAGTGTCTTCTTCACTT 28

RESULT 9
US-09-960-352-11316
Sequence 11316, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 11316
LENGTH: 387
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 48-LIB34-079-Q1-E1-D8
US-09-960-352-11316

Query Match 4.8%; Score 180.6; DB 10; Length 387;
Best Local Similarity 78.5%; Pred. No. 1.2e-47;
Matches 216; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

1479 CAGAGCGCGGCTGTGTGCTGCTGACCTTGAAGTCAGAGAGCCCAAGCCAGCTGACAGAGC 1538
111 CAGGCCAGAGCTGTGAGAGCTGTGACCTTGAAGTCTGAAGAGCTCAAGCCAGCTGACAGAGC 170
1539 TTCCTTGGCCCGCATGAGAGGTAGAGAGGCTTGTATGGAAGATCCCTTGGCCCTTA 1598
171 TTCCTTGGCTCCATACGAAAGTGTGATGGGCTCTGTATGAGAGATCTCTGGCCCCC 230
1599 ACTCATGTTCGGCAATGTCCCGCTTGTACTCATCTCATGTGTCCCAATTTGTACTGCT 1658
231 ACTCGTATGTGCAATGTCTGGGCTTGTACCTCGTCAAGGCCCAATCTGTACTGCT 290
1659 ACAGTATGTGAGAACTGATGACAGCCAGGCCCTTCCGCAACCCCAAGATGTGACCA 1718

291 ACAAGATGTGAGAGATTGAGCAAGACCAAGGACCTTCCCGAGGCCCAAGAGATGAACA 350
1719 GTCTCTGGGCTGTGACTCACTGAGGGGCTTCTGTA 1753
351 GCCTCTGGGCTGTGCTCCCAAGCAGCACTCTGTA 385

RESULT 10
US-09-960-352-7040
Sequence 7040, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 7040
LENGTH: 408
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 30-LIB34-026-Q1-E1-H5
US-09-960-352-7040

Query Match 4.5%; Score 167; DB 10; Length 408;
Best Local Similarity 79.8%; Pred. No. 3.5e-43;
Matches 197; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

1479 CAGAGCGCGGCTGTGTGCTGCTGACCTTGAAGTCAGAGAGCCCAAGCCAGCTGACAGAGC 1538
162 CAGGCCAGAGCTGTGAGAGCTGTGACCTTGAAGTCTGAAGAGCTCAAGCCAGCTGACAGAGC 221
1539 TTCCTTGGCCCGCATGAGAGGTAGAGAGGCTTGTATGGAAGATCCCTTGGCCCTTA 1598
222 TTCCTTGGCTCCATACGAAAGTGTGATGGGCTCTGTATGAGAGATCTCTGGCCCCC 281
1599 ACTCATGTTCGGCAATGTCCCGCTTGTACTCATCTCATGTGTCCCAATTTGTACTGCT 1658
282 ACTCGTATGTGCAATGTCTGGGCTTGTACCTCGTCAAGGCCCAATCTGTACTGCT 341
1659 ACAGTATGTGAGAACTGATGACAGCCAGGCCCTTCCGCAACCCCAAGATGTGACCA 1718
342 ACAAGATGTGAGAGATTGAGCAAGACCAAGGACCTTCCCGAGGCCCAAGAGATGAACA 401
1719 GTCTCTCT 1725
402 GCCTCTCT 408

RESULT 11
US-09-918-995-30754
Sequence 30754, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyeeg, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: PaedSeq for Windows Version 3.0
SEQ ID NO 30754
LENGTH: 488
TYPE: DNA

Db 60 TGAGTCTG--CTGNCACCTCCAGTTTCATATAAAGACAGATTGCANTTCAC 3
 Qy 1899 TT 1900
 Db 2 TT 1

RESULT 6 US-10-316-253-39

Sequence 39, Application US/10316253
 Publication No. US20030162706A1
 GENERAL INFORMATION:
 APPLICANT: The Procter & Gamble Company
 APPLICANT: Peters, Kevin
 APPLICANT: Thompson, Larry
 APPLICANT: Wang, Feng
 APPLICANT: Greib, Kenneth
 TITLE OF INVENTION: Angiogenesis Modulating Proteins
 FILE REFERENCE: 8865M
 CURRENT APPLICATION NUMBER: US/10/316,253
 CURRENT FILING DATE: 2002-12-10
 PRIOR FILING DATE: 2002-02-08
 PRIOR APPLICATION NUMBER: US 60/355,295
 NUMBER OF SEQ ID NOS: 308
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 39
 LENGTH: 1516
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (55)..(1437)
 OTHER INFORMATION:
 US-10-316-253-39

Query Match 5.3%; Score 198.6; DB 13; Length 1516;
 Best Local Similarity 75.1%; Pred. No. 4.2e-53;
 Matches 262; Conservative 0; Mismatches 84; Indels 3; Gaps 1;

Qy 1471 CTTCTCCCGAGAGCGGGGTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAGTG 1530
 Db 1167 CGTCAATCAGAGCGGGGCTTTGGCTGAGACCTGAAGTCAGAGCCCAAGCCAGTG 1226
 Qy 1531 GACAGAGCTTCTTGCGCCCATGAGAGGTAGACGAGCCCTTGATGAGAAAGTCCCT 1590
 Db 1227 GGCAGAGCTTCTTGCGCCCATGAGAGGTAGAGTGGTGGCTTGTGAGAAAGTCCCT 1286
 Qy 1591 TGGCCCTAATCTAGTTCGCGCAATGCTCCGCTGTATCTCATCTCATGTCCTCAATT 1650
 Db 1287 TGTCTCTACTAGTGTCTTCCATAGTGTCCCACTTGTCTTATCATGAGCCCAATT 1346
 Qy 1651 GTACTGCTACAGTGTGTGAGAACTGAATGACGCCAAGCCCTTCCGCAATCCAGAA 1710
 Db 1347 ATACTGCTATAGAGATATAGCAAACTGAATGACGCCAAGTCTGCTCAGCCCGAGAA 1406
 Qy 1711 TGTGACAGTCTCTGGGCTGCACTGAGGGGCC--TTCTGACATGAGTGGCCT 1767
 Db 1407 AGTGAACAGATCTTGGCTGAGTCAATATAAAGCCCTGATGAGAAATTAAGCCAGCCA 1466
 Qy 1768 GGGCCCACTCTAGTTCCTCATATATAAGACAGATGCTTTCGCTT 1816
 Db 1467 CCCCACTCTCCATTTCATTCTAATAAACCAGATGTTTCTTCACAT 1515

RESULT 7
 US-10-175-523-84
 Sequence 84, Application US/10175523
 Publication No. US20030096264A1
 GENERAL INFORMATION:
 APPLICANT: Brockman, Jeffrey
 APPLICANT: Evans, David
 APPLICANT: Hook, Derek
 APPLICANT: Klimczak, Leszek

APPLICANT: Laeng, Pascal
 APPLICANT: Palfreyman, Michael
 APPLICANT: Rajan, Prithi
 TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
 FILE REFERENCE: 3235/10795-US3
 CURRENT APPLICATION NUMBER: US/10/175,523
 CURRENT FILING DATE: 2002-06-18
 PRIOR APPLICATION NUMBER: US 60/299,151
 PRIOR FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: US 60/317,828
 PRIOR FILING DATE: 2001-09-07
 PRIOR APPLICATION NUMBER: US 60/325,150
 PRIOR FILING DATE: 2001-09-25
 PRIOR APPLICATION NUMBER: US 60/333,047
 PRIOR FILING DATE: 2001-11-14
 PRIOR APPLICATION NUMBER: US 60/349,936
 PRIOR FILING DATE: 2002-01-18
 PRIOR APPLICATION NUMBER: US 60/361,834
 PRIOR FILING DATE: 2002-03-04
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 84
 LENGTH: 1516
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 US-10-175-523-84

Query Match 5.3%; Score 198.6; DB 15; Length 1516;
 Best Local Similarity 75.1%; Pred. No. 4.2e-53;
 Matches 262; Conservative 0; Mismatches 84; Indels 3; Gaps 1;

Qy 1471 CTTCTCCCGAGAGCGGGGTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAGTG 1530
 Db 1167 CGTCAATCAGAGCGGGGCTTTGGCTGAGACCTGAAGTCAGAGCCCAAGCCAGTG 1226
 Qy 1531 GACAGAGCTTCTTGCGCCCATGAGAGGTAGACGAGCCCTTGATGAGAAAGTCCCT 1590
 Db 1227 GGCAGAGCTTCTTGCGCCCATGAGAGGTAGAGTGGTGGCTTGTGAGAAAGTCCCT 1286
 Qy 1591 TGGCCCTAATCTAGTTCGCGCAATGCTCCGCTGTATCTCATCTCATGTCCTCAATT 1650
 Db 1287 TGTCTCTACTAGTGTCTTCCATAGTGTCCCACTTGTCTTATCATGAGCCCAATT 1346
 Qy 1651 GTACTGCTACAGTGTGTGAGAACTGAATGACGCCAAGCCCTTCCGCAATCCAGAA 1710
 Db 1347 ATACTGCTATAGAGATATAGCAAACTGAATGACGCCAAGTCTGCTCAGCCCGAGAA 1406
 Qy 1711 TGTGACAGTCTCTGGGCTGCACTGAGGGGCC--TTCTGACATGAGTGGCCT 1767
 Db 1407 AGTGAACAGATCTTGGCTGAGTCAATATAAAGCCCTGATGAGAAATTAAGCCAGCCA 1466
 Qy 1768 GGGCCCACTCTAGTTCCTCATATATAAGACAGATGCTTTCGCTT 1816
 Db 1467 CCCCACTCTCCATTTCATTCTAATAAACCAGATGTTTCTTCACAT 1515

RESULT 8
 US-09-960-352-6580/c
 Sequence 6580, Application US/09960352
 Patent No. US20020137139A1
 GENERAL INFORMATION:
 APPLICANT: Warren, Wesley C.
 APPLICANT: Tao, Ningbing
 APPLICANT: Byatt, John C.
 APPLICANT: Mathialagan, Nagappan
 TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 FILE REFERENCE: 16511.006/37-21(10298)C
 CURRENT APPLICATION NUMBER: US/09/960,352
 CURRENT FILING DATE: 2001-09-24
 NUMBER OF SEQ ID NOS: 15112
 SEQ ID NO 6580
 LENGTH: 400

QY 3568 AGTTAGGCTGCTATAGAGAAATATCTTAGAGGGGATATCTCAGCAATAGGAATTTA 3627
DB 3017 AGTTAGGCTGCTATAGAGAAATATCTTAGAGGGGATATCTCAGCAATAGGAATTTA 3076
QY 3628 TTGTTTCACATTTCTGAGGCTGGAATAATCCAGATCAAGGCTCCAGAGGTTCAAGTCTC 3687
DB 3077 TTGTTTCACATTTCTGAGGCTGGAATAATCCAGATCAAGGCTCCAGAGGTTCAAGTCTC 3136
QY 3688 GCTGAGGCTGTTTCTGCTCTGGAAGATGCACTTTTGTGTTCTCA 3737
DB 3137 GCTGAGGCTGTTTCTGCTCTGGAAGATGCACTTTTGTGTTCTCA 3186

RESULT 3

US-10-125-237-19
; Sequence 19, Application US/10125237
; Publication No. US20030022329A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dermanac, Radoje T.
; TITLE OF INVENTION: No. US20030022329A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/125,237
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 19
; LENGTH: 1631
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (71)..(1459)
US-10-125-237-19

Query Match 11.9%; Score 446; DB 15; Length 1631;
Best Local Similarity 73.7%; Pred. No. 4,7e-134;
Matches 722; Conservative 0; Mismatches 5; Indels 253; Gaps 2;

QY 928 TCCACAGGAGCCCACTACTGCGCTGAGACACAGCCGGAGTGGCTGACATAGCTGGCCC 987
DB 899 TTCAAGTGGAGCCCACTACTGCGCTGAGACACAGCCGGAGTGGCTGACATAGCTGGCCC 958
QY 988 ATTGCTCATGAGTGGCCCAAGGGTCTTCAAGCAAGTGAATCTGCTTTCTGGGAAGA 1047
DB 959 ATTGCTCATGAGTGGCCCAAGGGTCTTCAAGCAAGTGAATCTGCTTTCTGGGAAGA 1018
QY 1048 AAACCTCATCTGCTCCAGGTGTGTATTTGGGGAGAGGCTTAGAGTAGACTGGGACAG 1107
DB 1019 AAACCTCATCTGCTCCAGGTGTGTATTTGGGGAGAGGCTTAGAGTAGACTGGGACAG 1031
QY 1108 CATATCAACTCTGTATTTATTAACATCTTTGCTCCAGGGACCCAGATATATGCTC 1167
DB 1032 -----TCAGGGGACCCAGATATATGCTC 1055
QY 1168 TCCTGACAAAGGAGGCTATACCTTAGTAGACGTTATCCGAAGCGGCTGGAGAAGAG 1227
DB 1056 TCCTGACAAAGGAGGCTATACCTTAGTAGACGTTATCCGAAGCGGCTGGAGAAGAG 1115

QY 1228 TGGGAGCCCTCATAGGATTAATCTGGAATCTGTTGATGCGGCTTTATCTGCGCTGGGT 1287
DB 1116 TGGGAGCCCTCATAGGATTAATCTGGAATCTGTTGATGCGGCTTTATCTGCGCTGGGT 1175
QY 1288 CTTCCTGGCTCCATATCATATGAGAGGTGAGGGGCTTCTGGGTCTTAGAGGGACCTTGT 1347
DB 1176 CTTCCTGGCTCCATATCATATGAGAGGTGAGGGGCTTCTGGGTCTTAGAGGGACCTTGT 1196
QY 1348 CTGCTACTGTCTGTGGCATTAATCCCAACAGGGCATAGAAAGGCTTAGGTACAGATCC 1407
DB 1197 ----- 1196
QY 1408 CCAGGGCATGAGAGGCTTAGGTACAGATCCCATGATGAGAGGCCATGTATGTTGG 1467
DB 1197 ----- 1196
QY 1468 TGGCTTCTCCCAAGAGCGGCTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAC 1527
DB 1197 -----CAGAGCGGCTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAC 1245
QY 1528 GTGAGACAGGCTTCTTGGCCCCCATAGAAAGGTAGACGAGGCTTGTATAGAAAGTC 1587
DB 1246 GTGAGACAGGCTTCTTGGCCCCCATAGAAAGGTAGACGAGGCTTGTATAGAAAGTC 1305
QY 1588 CCTTGGCCCTTAATCATATGTTCCGCCCAATGATGCCGCTTGTACCTCATATGATGCCAA 1647
DB 1306 CCTTGGCCCTTAATCATATGTTCCGCCCAATGATGCCGCTTGTACCTCATATGATGCCAA 1365
QY 1648 TTGTATCTGCTACAGTATGTAGAGAACTGAATGAGCCCAAGGCCCTTCCGCAACCCCA 1707
DB 1366 TTGTATCTGCTACAGTATGTAGAGAACTGAATGAGCCCAAGGCCCTTCCGCAACCCCA 1425
QY 1708 GAATGTACCAATCTTCTGCGCTGACATCACTGAGGGGCTTCTGACATGAGTCTGGCCT 1767
DB 1426 GAATGTACCAATCTTCTGCGCTGACATCACTGAGGGGCTTCTGACATGAGTCTGGCCT 1485
QY 1768 GGGCCCACTCTAGTCTCTCAATTAATTAAGACAGATGCTTCTGCTGCTCACTGAGGG 1827
DB 1486 GGGCCCACTCTAGTCTCTCAATTAATTAAGACAGATGCTTCTGCTGCTCACTGAGGG 1545
QY 1828 GCTTCTGACATGATGTGAGCTGAGCCCACTCCCACTTCTCATTAATAAGACAGAT 1887
DB 1546 GCTTCTGACATGATGTGAGCTGAGCCCACTCCCACTTCTCATTAATAAGACAGAT 1605
QY 1888 TGCTTCTTCACTTGAATCAA 1907
DB 1606 TGCTTCTTCACTTGAATCAA 1625

RESULT 4

US-10-105-891-19
; Sequence 19, Application US/10105891
; Publication No. US20030073099A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dermanac, Radoje T.
; TITLE OF INVENTION: No. US20030073099A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2A
; CURRENT APPLICATION NUMBER: US/10/105,891
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22

Db 928 ----- 927
Qy 1408 CCAGGCGATGAGAGGCTTAGGTCCAGATCCCATGACATGAAAGCCATGTTGG 1467
Db 928 ----- 927
Qy 1468 TGCCTTCTCCCGAGACGGGGCTGTGGTGGCTTGACCTTGAAGTCAGAGAGCCCAAGCCAC 1527
Db 928 -----CAGAGCGCGGCTGTGGTGGCTTGACCTTGAAGTCAGAGAGCCCAAGCCAC 976
Qy 1528 GTGAGACAGAGCTTCTTGGGCCCATGAGAGTAGACGAGACCTTGTGTGAAAAATGTC 1587
Db 977 GTGAGACAGAGCTTCTTGGGCCCATGAGAGTAGACGAGACCTTGTGTGAAAAATGTC 1036
Qy 1588 CTTGGCCCTTAAGTCATGTTCGCGCAATGGTCCCGGCTTGTACCTCATCATGTGTCCTCA 1647
Db 1037 CTTGGCCCTTAAGTCATGTTCGCGCAATGGTCCCGGCTTGTACCTCATCATGTGTCCTCA 1096
Qy 1648 TTTGTACTGTACAGATGATGTGAGAACTGAAATGACGCCAAGGCCCTTCCGAAACCCCA 1707
Db 1097 TTTGTACTGTACAGATGATGTGAGAACTGAAATGACGCCAAGGCCCTTCCGAAACCCCA 1156
Qy 1708 GAATGACACAGTCTCTGGGCTGCACTCACTGAGGGGCCCTTGTGACATGATCTGGCCT 1767
Db 1157 GAATGACACAGTCTCTGGGCTGCACTCACTGAGGGGCCCTTGTGACATGATCTGGCCT 1216
Qy 1768 GGGCCCACTCTAGTTCCTCATATATAGACAGATTTGCTTTCGCTTCACTGAGGG 1827
Db 1217 GGGCCCACTCTAGTTCCTCATATATAGACAGATTTGCTTTCGCTTCACTGAGGG 1276
Qy 1828 GCTTTCTGACATAGTGTGGCTGGCCCACTCCCAAGTTTCTCAATATATAGACAGAT 1887
Db 1277 GCTTTCTGACATAGTGTGGCTGGCCCACTCCCAAGTTTCTCAATATATAGACAGAT 1336
Qy 1888 TGCCTTCTTCACTGAATCAAGGAGCTTGTGCGAAACAATCTTCTTCTTGAAGTTGA 1947
Db 1337 TGCCTTCTTCACTGAATCAAGGAGCTTGTGCGAAACAATCTTCTTCTTGAAGTTGA 1396
Qy 1948 AAGATTGACATTTCTCTTTGAGGGTGTGCAAGCTCAAAACAAGCTGTGAGAAACAAGGA 2007
Db 1397 AAGATTGACATTTCTCTTTGAGGGTGTGCAAGCTCAAAACAAGCTGTGAGAAACAAGGA 1456
Qy 2008 GGGGAGACCTAAGGGGGAACCTATCTGTGCGAGATTTCTTGAAGTCCAGATCATAA 2067
Db 1457 GGGGAGACCTAAGGGGGAACCTATCTGTGCGAGATTTCTTGAAGTCCAGATCATAA 1516
Qy 2068 CTAGCTCTTGCAGACTATCTACACATAGTGGGGGGAAGAAACAGATCGGAAGAGG 2127
Db 1517 CTAGCTCTTGCAGACTATCTACACATAGTGGGGGGAAGAAACAGATCGGAAGAGG 1576
Qy 2128 AACAGCTGAGTTTATACAGCAAGTAGAGGTGAGCTAGGACTCTGTATTCACCTTGTGG 2187
Db 1577 AACAGCTGAGTTTATACAGCAAGTAGAGGTGAGCTAGGACTCTGTATTCACCTTGTGG 1636
Qy 2188 TAATATGCCCAACCCGCGGCAAGGATCAAGAAACAAGGGCTTGGGGCACTATGCA 2247
Db 1637 TAATATGCCCAACCCGCGGCAAGGATCAAGAAACAAGGGCTTGGGGCACTATGCA 1696
Qy 2248 TGTGCAAAAGAGATTGGCTCAGAGTTGTGGGTGAGAGTCAATCTGGGGGCACTCAAA 2307
Db 1697 TGTGCAAAAGAGATTGGCTCAGAGTTGTGGGTGAGAGTCAATCTGGGGGCACTCAAA 1756
Qy 2308 TTAATGTTCTGGGTGATTCAAGTACACACTCATGGCTTGTGTTCATGATGATTAGCA 2367
Db 1757 TTAATGTTCTGGGTGATTCAAGTACACACTCATGGCTTGTGTTCATGATGATTAGCA 1816
Qy 2368 TGAACAATGGAATGAATGAAGTGGGGAACAGAAATACACAGCTGTGTGTAGAGGC 2427
Db 1817 TGAACAATGGAATGAATGAAGTGGGGAACAGAAATACACAGCTGTGTGTAGAGGC 1876
Qy 2428 AAGCTGAGAGAGAGAAAGAAATGAATGGCACATGAGACATTTGACAGAACAGT 2487

Db 1877 AAGCTGAGAGAGAGAAAGAAATGAATGGCACATGAGACATTTGACAGAACAGT 1936
Qy 2488 CCTGGGAGTCTTGTGAGAGGCTCAGAGACTTGTCTGGGACAGAGGATCGGCTACCA 2547
Db 1937 CCTGGGAGTCTTGTGAGAGGCTCAGAGACTTGTCTGGGACAGAGGATCGGCTACCA 1996
Qy 2548 ATTAGGCTCTGGGATCTGCAACATCTAGACAGAGAAATGTCACTGGGACAGAGGTTGC 2607
Db 1997 ATTAGGCTCTGGGATCTGCAACATCTAGACAGAGAAATGTCACTGGGACAGAGGTTGC 2056
Qy 2608 AGTCTGTGTTATTTTCTGGTCAATAAAGTGAAGGCTTGGTTCCAAATTTGCTGAC 2667
Db 2057 AGTCTGTGTTATTTTCTGGTCAATAAAGTGAAGGCTTGGTTCCAAATTTGCTGAC 2116
Qy 2668 AGTCTTTTAATGGGAATTTGGGCTTGAATTAAGTATGCTTCAAGAGGCCCTA 2727
Db 2117 AGTCTTTTAATGGGAATTTGGGCTTGAATTAAGTATGCTTCAAGAGGCCCTA 2176
Qy 2728 TTTCTCACTGCTTTTACAACCCAGCTGAGGTTGAGGCTGCTTGTTCAGGCTCAAAA 2787
Db 2177 TTTCTCACTGCTTTTACAACCCAGCTGAGGTTGAGGCTGCTTGTTCAGGCTCAAAA 2236
Qy 2788 AATAGCTGAGTTTCCAGAGAGGGCTTATTTCTGAGCTTCGTTGCTTACCTCATTT 2847
Db 2237 AATAGCTGAGTTTCCAGAGAGGGCTTATTTCTGAGCTTCGTTGCTTACCTCATTT 2296
Qy 2848 TCTTTTCTGTAAATATAGACAAATGCCACCTTCCAGTACAAATGATATTAATC 2907
Db 2297 TCTTTTCTGTAAATATAGACAAATGCCACCTTCCAGTACAAATGATATTAATC 2356
Qy 2908 AAACCCATCCCTTGAACCTGCTTGGGAGGGGCTGTGAGTGAACCCAGACTGTGGCTC 2967
Db 2357 AAACCCATCCCTTGAACCTGCTTGGGAGGGGCTGTGAGTGAACCCAGACTGTGGCTC 2416
Qy 2968 ATGGCTCATATGATATGAGATGAGTCAAGCCCTCCCAACTGTGACCAATTTGCTCGTAGA 3027
Db 2417 ATGGCTCATATGATATGAGATGAGTCAAGCCCTCCCAACTGTGACCAATTTGCTCGTAGA 2476
Qy 3028 CTTTGAATGGGTAGATGATAGTAAACAAGCTCTGATCAACAAGGCTTGTACTGGG 3087
Db 2477 CTTTGAATGGGTAGATGATAGTAAACAAGCTCTGATCAACAAGGCTTGTACTGGG 2536
Qy 3088 AGGCAAGCTATATGATGAGTCTCAAGCTTAAAGGGCTGGGAGCTGGGGGACAGTGTCTAG 3147
Db 2537 AGGCAAGCTATATGATGAGTCTCAAGCTTAAAGGGCTGGGAGCTGGGGGACAGTGTCTAG 2596
Qy 3148 ATTAGGCTTAACTAGAAATTTGACTGAGCTGAGAAACAGAGGTTAAGGGCCAAAGACA 3207
Db 2597 ATTAGGCTTAACTAGAAATTTGACTGAGCTGAGAAACAGAGGTTAAGGGCCAAAGACA 2656
Qy 3208 GGGTGTGGGTCTACTCTCTTGAAGACCTTGAAGCTTTTACTTTTCAATTTGTGTTC 3267
Db 2657 GGGTGTGGGTCTACTCTCTTGAAGACCTTGAAGCTTTTACTTTTCAATTTGTGTTC 2716
Qy 3268 TTGGATGGGCTAACCTCAGCGGGGTTGGCTGACTTAAAGGGGTGAGAACAGAGACAGAT 3327
Db 2717 TTGGATGGGCTAACCTCAGCGGGGTTGGCTGACTTAAAGGGGTGAGAACAGAGACAGAT 2776
Qy 3328 TTGAGTCTGGTCTTATCAAGTTTCACTGACCTTGGGACCACTGCTGATCATGCC 3387
Db 2777 TTGAGTCTGGTCTTATCAAGTTTCACTGACCTTGGGACCACTGCTGATCATGCC 2836
Qy 3388 AGGAGCTTGAAGGTGTCTTAAACAATTTTCAACAATGGAATACCAAGGTTAACTTTC 3447
Db 2837 AGGAGCTTGAAGGTGTCTTAAACAATTTTCAACAATGGAATACCAAGGTTAACTTTC 2896
Qy 3448 TCTTGTTTTCAAGAGGAGGAGTAAAGTCTCCCTTCTCTCTTCCACAGTGTTC 3507
Db 2897 TCTTGTTTTCAAGAGGAGGAGTAAAGTCTCCCTTCTCTCTTCCACAGTGTTC 2956
Qy 3508 TCTTGAAGGGAATCTCTAGCTTGTCTCAAGGAACTCCCAAGAAATGTTTGTTCAGTC 3567
Db 2957 TCTTGAAGGGAATCTCTAGCTTGTCTCAAGGAACTCCCAAGAAATGTTTGTTCAGTC 3016

QY 2401 GAAATACACAGCTGTGTCTCAAGGCAAGCTGAGAGAGAGAGAGAGAGATGATGCA 2460
 Db 12401 GAAATACACAGCTGTGTCTCAAGGCAAGCTGAGAGAGAGAGAGAGATGATGCA 12460
 QY 2461 CCATGAGCAATTTTGAGAGACAGTCTCTGGAGTCTTGTGAGAGCTTCAAGAGCTTT 2520
 Db 12461 CCATGAGCAATTTTGAGAGACAGTCTCTGGAGTCTTGTGAGAGCTTCAAGAGCTTT 12520
 QY 2521 GCTGGACAGAGGATCTGAGCTACCCCAATTAGCTCTGGAGTCTTGTGAGAGCTTCAAG 2580
 Db 12521 GCTGGACAGAGGATCTGAGCTACCCCAATTAGCTCTGGAGTCTTGTGAGAGCTTCAAG 12580
 QY 2581 AGCAAAATGTCATGAGAGAGAGGTTGAGTCTTGTGATTTTCTGTGATTAATCTGAT 2640
 Db 12581 AGCAAAATGTCATGAGAGAGAGGTTGAGTCTTGTGATTTTCTGTGATTAATCTGAT 12640
 QY 2641 AAGGCTTTGGGTTCCAAATTTGCTGAGAGCTTGTAACTGGAAATTTGGGCTTGA 2700
 Db 12641 AAGGCTTTGGGTTCCAAATTTGCTGAGAGCTTGTAACTGGAAATTTGGGCTTGA 12700
 QY 2701 GGTAGCTATGTCAGAGAGAGGCTTATCTCACTGAGCTTAAACAACGAGTGAAGTT 2760
 Db 12701 GGTAGCTATGTCAGAGAGAGGCTTATCTCACTGAGCTTAAACAACGAGTGAAGTT 12760
 QY 2761 GAGAGCTGAGTGTTCAGGCTCAAAAATAATAGCTGATTTCCAGAGAGGCTTAT 2820
 Db 12761 GAGAGCTGAGTGTTCAGGCTCAAAAATAATAGCTGATTTCCAGAGAGGCTTAT 12820
 QY 2821 CTGAGCTTCCGTCTCTAGCTCAATTTCTTTCTCTGTAAATAGACATGCAACCA 2880
 Db 12821 CTGAGCTTCCGTCTCTAGCTCAATTTCTTTCTCTGTAAATAGACATGCAACCA 12880
 QY 2881 CCTTCCAGTGAACAATGATATAGACTCAACCAATCCCTTGAACCTGTTGGAGAGG 2940
 Db 12881 CCTTCCAGTGAACAATGATATAGACTCAACCAATCCCTTGAACCTGTTGGAGAGG 12940
 QY 2941 TCTGAGAGTGAAGCCAGAGCTGCTCATGCTCATGATGATGAGAGTCAAGCCCTCC 3000
 Db 12941 TCTGAGAGTGAAGCCAGAGCTGCTCATGCTCATGATGATGAGAGTCAAGCCCTCC 13000
 QY 3001 AACCTGTCAAGCTTTGCTCTGTAAGACTTTGATGAGTGAAGTGAAGTGAAGCTCT 3060
 Db 13001 AACCTGTCAAGCTTTGCTCTGTAAGACTTTGATGAGTGAAGTGAAGTGAAGCTCT 13060
 QY 3061 GACTGTCAACAAGAGGCTTTGTACTGGAGAGGCTGATGAGTGTCTCAAGTTAAAG 3120
 Db 13061 GACTGTCAACAAGAGGCTTTGTACTGGAGAGGCTGATGAGTGTCTCAAGTTAAAG 13120
 QY 3121 GCTGGAGCTGGGGGACAGTGTCTCAGATTAGGGTCTAACTAGAGAGTGAAGTGAAG 3180
 Db 13121 GCTGGAGCTGGGGGACAGTGTCTCAGATTAGGGTCTAACTAGAGAGTGAAGTGAAG 13180
 QY 3181 AGAACAAGAGTGAAGGGGCCAGAGAGAGGTTGTGGTCTAATCTTTAGAGAGCACTTGA 3240
 Db 13181 AGAACAAGAGTGAAGGGGCCAGAGAGAGGTTGTGGTCTAATCTTTAGAGAGCACTTGA 13240
 QY 3241 GCTTTACTTTTATCTATATGAGTCTTGGATGAGTCAACCTCAAGGGGTGGCTGAG 3300
 Db 13241 GCTTTACTTTTATCTATATGAGTCTTGGATGAGTCAACCTCAAGGGGTGGCTGAG 13300
 QY 3301 TCTAAGGGGTGAGAGAGAGAGAGTTCAGGTCGTGCTCTTATCAAGTTCAAGTCA 3360
 Db 13301 TCTAAGGGGTGAGAGAGAGAGAGTTCAGGTCGTGCTCTTATCAAGTTCAAGTCA 13360
 QY 3361 CACTTGGAGCACTGCTGATCATGCAAGGAGCTTGAAGTGTCTAAACAAGTTATCCA 3420
 Db 13361 CACTTGGAGCACTGCTGATCATGCAAGGAGCTTGAAGTGTCTAAACAAGTTATCCA 13420
 QY 3421 CAACGTGTATCCCAAGGTTAACTTTCTCTGTTTCAAGAGGAGAGATCAAGTCTC 3480
 Db 13421 CAACGTGTATCCCAAGGTTAACTTTCTCTGTTTCAAGAGGAGAGATCAAGTCTC 13480
 QY 3481 CCTTTCTCTTTCTCTCAAGGTTCTCTGAGAGGAAATCTCTAGCTTGTCTCAAGG 3540

Db 13481 CCTTTCTCTTTCTCTCCACAGTGTCTCTGAGAGGAAATCTCTAGCTTGTCTCCAGG 13540
 QY 3541 AACTCCAGAAATGTTTGTTCAGTCAAGTTTATGAGTGTCTATTAAGAAATATCTTGA 3600
 Db 13541 AACTCCAGAAATGTTTGTTCAGTCAAGTTTATGAGTGTCTATTAAGAAATATCTTGA 13600
 QY 3601 GGGTAATCTATCAGCAATAGAAATTAATGTTCAAAATCTGAGAGCTGGAATAATCAAG 3660
 Db 13601 GGGTAATCTATCAGCAATAGAAATTAATGTTCAAAATCTGAGAGCTGGAATAATCAAG 13660
 QY 3661 ATCAAGGCTTCAGAGAGTTCAGTCTGAGTGTGAGTGTCTTGTGGAAGATGAGCACC 3720
 Db 13661 ATCAAGGCTTCAGAGAGTTCAGTCTGAGTGTGAGTGTCTTGTGGAAGATGAGCACC 13720
 QY 3721 TTTTGTCTGTCTCTCA 3737
 Db 13721 TTTTGTCTGTGTCTCA 13737

RESULT 2

US-09-900-448-1
 ; Sequence 1, Application US/0900448
 ; Publication No. US20030220488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CECARDI, Toni et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
 ; FILE REFERENCE: CL001272
 ; CURRENT APPLICATION NUMBER: US/09/900,448
 ; CURRENT FILING DATE: 2001-07-09
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3186
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-900-448-1

Query Match 61.0%; Score 2279.2; DB 13; Length 3186;
 Best Local Similarity 90.9%; Pred. No. 0;
 Matches 2554; Conservative 0; Mismatches 3; Indels 253; Gaps 2;
 QY 928 TCACAGAGGAGCCCACTATCTGAGTCTGAGACACAGCCGGAGAGTGTGATAGTCTGCC 987
 Db 630 TTCAGTGGAGCCCACTATCTGAGTCTGAGACACAGCCGGAGAGTGTGATAGTCTGCC 689
 QY 988 ATTGCTCATCATGAGGCCCAAGGAGTCTTCAAGAGTGAAGTGTCTTCTTCTGGAGAA 1047
 Db 690 ATTGCTCATCATGAGGCCCAAGGAGTCTTCAAGAGTGAAGTGTCTTCTTCTGGAGAA 749
 QY 1048 AAATCTATCTGATCAGAGTGTGATTTGGGGAGAGGCTTGAAGTGAAGACTGGAGCAAG 1107
 Db 750 AAATCTATCTG----- 762
 QY 1108 CATATCCAACTGTATTTATTTATACATCTTTGTCTCCAGGGCAACCAAGTATATGCT 1167
 Db 763 -----TCCAGGGCAACCAAGTATATGCT 786
 QY 1168 TCTGTACAAAGGAGGAGTATACCTAGTGAAGGTTATCCGAAGGAGGCTGAGAGAGAG 1227
 Db 787 TCTGTACAAAGGAGGAGTATACCTAGTGAAGGTTATCCGAAGGAGGCTGAGAGAGAG 846
 QY 1228 TGGGAGCCCTCATGAGATTAATCTGAGACTGTGAGATGCGGCTTTATCTGCTGGGT 1287
 Db 847 TGGGAGCCCTCATGAGATTAATCTGAGACTGTGAGATGCGGCTTTATCTGCTGGGT 906
 QY 1288 CTTCCTGGGTCATATCATGAGGAGGAGGCTTCTGGGTCTTGAAGGAGGAGCTTGT 1347
 Db 907 CTTCCTGGGTCATATCATG----- 927
 QY 1348 CTGCTACCTGTGTGAGCATATATCCCAACAGGGAGATGAGAGGCTTATGTCAGAGATCC 1407

Db 10181 AAAAAGATAGAGAAATCTGGAGAGTCAAGGTGTACAGAAAGCCAGTTCCAAAAA 10240
Qy 241 GACATTTAAAGGAAAGAGTAGTGAAGCTGCTCCTCGAAGAGTGAAGGTCAAGT 300
Db 10241 GACATTTAAAGGAAAGAGTAGTGAAGCTGCTCCTCGAAGAGTGAAGGTCAAGT 10300
Qy 301 GAGAACAGAGAAATGACATGAGATTTGCAATTGAGAAATCTAGCAACTGATTAAG 360
Db 10301 GAGAACAGAGAAATGACATGAGATTTGCAATTGAGAAATCTAGCAACTGATTAAG 10360
Qy 361 AACCAATTTCAATGTTGAGAGGAAACAGAACTGTAAATTGAAGAGATTGAGAAAAAGAC 420
Db 10361 AACCAATTTCAATGTTGAGAGGAAACAGAACTGTAAATTGAAGAGATTGAGAAAAAGAC 10420
Qy 421 AAATGGAGCTTGATTAATTTCTTAATTAAGTTGTTGAAAAAGAGAAAGAAAAACGGG 480
Db 10421 AAATGGAGCTTGATTAATTTCTTAATTAAGTTGTTGAAAAAGAGAAAGAAAAACGGG 10480
Qy 481 GTCCTAGCCAGTACCTCCCTCACTCTCCACCACTCATAGGAGAGACTGGAGAACAC 540
Db 10481 GTCCTAGCCAGTACCTCCCTCACTCTCCACCACTCATAGGAGAGACTGGAGAACAC 10540
Qy 541 AGCAGAGTGAAGAACTTCAGTGAAGTGTGCTCTCTTTTAAATTGAGCACTGTAT 600
Db 10541 AGCAGAGTGAAGAACTTCAGTGAAGTGTGCTCTCTTTTAAATTGAGCACTGTAT 10600
Qy 601 TTCAATATCTAATACCGCACTCTCTGTCATGAGCACTGAAATCTTGAAGATGCCCG 660
Db 10601 TTCAATATCTAATACCGCACTCTCTGTCATGAGCACTGAAATCTTGAAGATGCCCG 10660
Qy 661 CAACCCATGATGTGGCTTACCTGAGAACTTAGCCATGTTTTCACACTTGGCTTTC 720
Db 10661 CAACCCATGATGTGGCTTACCTGAGAACTTAGCCATGTTTTCACACTTGGCTTTC 10720
Qy 721 TTTCAAGCACTGCTGATTCAGTTCAGCCAGGCAAGTGCACCAATGTGTCACCA 780
Db 10721 TTTCAAGCACTGCTGATTCAGTTCAGCCAGGCAAGTGCACCAATGTGTCACCA 10780
Qy 781 GTCTGTCTAATTTCTCTCTCTGACCTGGCTCTTCACTTGGCTCTGAGATTC 840
Db 10781 GTCTGTCTAATTTCTCTCTCTGACCTGGCTCTTCACTTGGCTCTGAGATTC 10840
Qy 841 CTCCTCTCACTGACTCATTTCTGATTCATCACTAGCTCTCTGCTGGCTTTCG 900
Db 10841 CTCCTCTCACTGACTCATTTCTGATTCATCACTAGCTCTCTGCTGGCTTTCG 10900
Qy 901 CAGCGGCTTGAAGCACTATGTATTCACAGGGAACCACTAGTGCCTTGAACAC 960
Db 10901 CAGCGGCTTGAAGCACTATGTATTCACAGGGAACCACTAGTGCCTTGAACAC 10960
Qy 961 AGCGGGATGCTGGCTAGCTAGCTGCTCATAGTGGCCCAAGGCTCTTACGCA 1020
Db 10961 AGCGGGATGCTGGCTAGCTAGCTGCTCATAGTGGCCCAAGGCTCTTACGCA 11020
Qy 1021 GTGAGTCTGCTTTCTGAGGAAAGAAAACTCATGTGCTGCAAGGTGATTAAGGGA 1080
Db 11021 GTGAGTCTGCTTTCTGAGGAAAGAAAACTCATGTGCTGCAAGGTGATTAAGGGA 11080
Qy 1081 GAGGCTTGAAGTAGAGCTGGGACAGACATATCCACTGTATTTAACAATCTTTG 1140
Db 11081 GAGGCTTGAAGTAGAGCTGGGACAGACATATCCACTGTATTTAACAATCTTTG 11440
Qy 1141 TCTCTCAGGGCACTCAGGTATATGTCTTCTGACAAAGAGGCTAATACCTAGTAAGC 1200
Db 11441 TCTCTCAGGGCACTCAGGTATATGTCTTCTGACAAAGAGGCTAATACCTAGTAAGC 11200
Qy 1201 GTTATCCGAAGGCTGAGAGAGAGAGTCCGCTCATAGGATTAATCTTGAAGCTG 1260
Db 11201 GTTATCCGAAGGCTGAGAGAGAGAGTCCGCTCATAGGATTAATCTTGAAGCTG 11260
Qy 1261 TGAATGCGGCTTATATCTGCTGGGTCTTCTCGGCTCATATCATGAGTGAAGGAGC 1320
Db 11261 TGAATGCGGCTTATATCTGCTGGGTCTTCTCGGCTCATATCATGAGTGAAGGAGC 11320

Qy 1321 TTCTGGGTCTTGAAGGAGCTTGTCTTCTACCTGTCTGTGGCATATCCCAACGAG 1380
Db 11321 TTCTGGGTCTTGAAGGAGCTTGTCTTCTACCTGTCTGTGGCATATCCCAACGAG 11380
Qy 1381 GGCATGAGAGGCTTGAAGTCAAGATCCCGAGGCAATGAGAGGCTTGAAGTCCCG 1440
Db 11381 GGCATGAGAGGCTTGAAGTCAAGATCCCGAGGCAATGAGAGGCTTGAAGTCCCG 11440
Qy 1441 ATGACATGAGAGCTATGTGTTGTGTGCTTCTTCCAGAGACGGCGCTGTGTGCT 1500
Db 11441 ATGACATGAGAGCTATGTGTTGTGTGCTTCTTCCAGAGACGGCGCTGTGTGCT 11500
Qy 1501 GGAACCTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
Db 11501 GGAACCTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11560
Qy 1561 AGACGAGGCTTGTATGAGAAAGTCCCTGAGGCTTAACTGATTTCCGCAATGCTC 1620
Db 11561 AGACGAGGCTTGTATGAGAAAGTCCCTGAGGCTTAACTGATTTCCGCAATGCTC 11620
Qy 1621 CGGCTTGAACCTGATCATGTGCTCCAAATTTGTAAGTGTGATGATGAGAACTGAA 1680
Db 11621 CGGCTTGAACCTGATCATGTGCTCCAAATTTGTAAGTGTGATGATGAGAACTGAA 11680
Qy 1681 TGAAGCAAGGAGGCTTCCGCAACCCGAGATGAGCACTCTCTGAGGCTGACCTG 1740
Db 11681 TGAAGCAAGGAGGCTTCCGCAACCCGAGATGAGCACTCTCTGAGGCTGACCTG 11740
Qy 1741 AGGAGGCTTCTGACATGATGCTGCTGAGCCCACTCTGATGCTCTCAATTAAGACA 1800
Db 11741 AGGAGGCTTCTGACATGATGCTGCTGAGCCCACTCTGATGCTCTCAATTAAGACA 11800
Qy 1801 GATTGCTTCTGCTTCTCACTGAGGAGGCTTCTGACATGATGCTGAGGCTGAGCCT 1860
Db 11801 GATTGCTTCTGCTTCTCACTGAGGAGGCTTCTGACATGATGCTGAGGCTGAGCCT 11860
Qy 1861 CCCAGTTCTGATTAATTAAGACAGATGCTCTTCACTGATCAAGGAGCCTTGTG 1920
Db 11861 CCCAGTTCTGATTAATTAAGACAGATGCTCTTCACTGATCAAGGAGCCTTGTG 11920
Qy 1921 TGAACAATCTTCTTCTTGAAGTGAAGAAAGTGAAGCACTTCTTGAAGGTGTGAGC 1980
Db 11921 TGAACAATCTTCTTCTTGAAGTGAAGAAAGTGAAGCACTTCTTGAAGGTGTGAGC 11980
Qy 1981 TCAACAGGCTGTGAGAAACAGGAGGAGAGCACTAAGGAGCAACCTATCTGCGC 2040
Db 11981 TCAACAGGCTGTGAGAAACAGGAGGAGAGCACTAAGGAGCAACCTATCTGCGC 12040
Qy 2041 AGATGATTTCTGAGTCAAGTCAATTAAGTCTTGAAGTCTTGAAGTCTTGAAGT 2100
Db 12041 AGATGATTTCTGAGTCAAGTCAATTAAGTCTTGAAGTCTTGAAGTCTTGAAGT 12100
Qy 2101 GGGAAAGAGAACAGAGTGAAGAGAAACAGTGAAGTATACAGAAATGAAGTGTG 2160
Db 12101 GGGAAAGAGAACAGAGTGAAGAGAAACAGTGAAGTATACAGAAATGAAGTGTG 12160
Qy 2161 AGCTAGAGCTTGAATCAACTGCTGTGATGAGGCAACACAGGCTGAGGATCAAG 2220
Db 12161 AGCTAGAGCTTGAATCAACTGCTGTGATGAGGCAACACAGGCTGAGGATCAAG 12220
Qy 2221 AACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Db 12221 AACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 12280
Qy 2281 AGAGAGTCAATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
Db 12281 AGAGAGTCAATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 12340
Qy 2341 ATGCTTGTGTTGCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2400
Db 12341 ATGCTTGTGTTGCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 12400

Result		Query				Description	
No.	Score	Match	Length	DB	ID		
1	3737	100.0	13737	13	US-09-900-448-3	Sequence 3,	App1
2	2279.2	61.0	3166	13	US-09-900-448-1	Sequence 1,	App1
3	446	11.9	1631	15	US-10-123-237-19	Sequence 19,	App1
4	446	11.9	1631	15	US-10-105-891-19	Sequence 19,	App1
5	265.2	7.1	504	13	US-09-880-107-3166	Sequence 3166,	App1
6	198.6	5.3	1516	10	US-10-316-253-39	Sequence 39,	App1
7	198.6	5.3	1516	15	US-10-175-523-94	Sequence 84,	App1
8	197.4	5.3	400	10	US-09-960-352-5368	Sequence 680,	App1
9	180.6	4.5	367	10	US-09-960-352-11316	Sequence 11316,	App1
10	167	4.8	468	10	US-09-960-352-7040	Sequence 7040,	App1
11	149.4	4.0	488	11	US-09-918-995-10754	Sequence 30754,	App1
12	134.2	3.6	493	11	US-09-918-995-32181	Sequence 32181,	App1
13	134.2	3.6	491	11	US-09-918-995-32820	Sequence 32820,	App1
14	123.6	3.3	420	10	US-09-960-352-10323	Sequence 10323,	App1
15	118.2	3.2	374	10	US-09-960-352-5368	Sequence 5368,	App1

16	117.4	3.1	383	10	US-09-960-352-6846	Sequence 6846, Ap
17	111.2	3.0	228	10	US-09-960-352-8240	Sequence 8240, Ap
18	106.4	2.8	27469	13	US-10-017-161-1015	Sequence 1015, Ap
19	106	2.8	357	10	US-09-960-352-12363	Sequence 12936, A
20	105.6	2.8	661	13	US-10-027-632-136875	Sequence 136875, S
21	105.6	2.8	661	14	US-10-027-632-136875	Sequence 136875, S
22	105.4	2.8	617	14	US-10-027-632-102912	Sequence 102912, S
23	105.4	2.8	617	14	US-10-027-633-102912	Sequence 102912, S
24	105.4	2.8	337658	10	US-09-813-320-3	Sequence 3, Appl
25	103.2	2.8	203	10	US-09-960-352-9190	Sequence 9190, Ap
26	102.4	2.7	599	13	US-10-027-632-236392	Sequence 236392, S
27	102.4	2.7	599	13	US-10-027-632-236393	Sequence 236393, S
28	102.4	2.7	599	13	US-10-027-633-236394	Sequence 236394, S
29	102.4	2.7	599	14	US-10-027-632-236392	Sequence 236392, S
30	102.4	2.7	599	14	US-10-027-632-236393	Sequence 236393, S
31	102.4	2.7	5599	14	US-10-027-632-236394	Sequence 236394, S
32	102.4	2.7	2522	13	US-10-027-632-235950	Sequence 235950, S
33	102.4	2.7	2532	14	US-10-027-633-255950	Sequence 255950, S
34	102.2	2.7	353	10	US-09-960-352-12972	Sequence 12972, A
35	101	2.7	715	13	US-10-027-632-202762	Sequence 202762, S
36	101	2.7	715	13	US-10-027-632-202763	Sequence 202763, S
37	101	2.7	715	14	US-10-027-632-202762	Sequence 202762, S
38	101	2.7	715	14	US-10-027-633-202763	Sequence 202763, S
39	100.2	2.7	593	9	US-09-864-761-11839	Sequence 11839, A
40	99.2	2.7	650	13	US-10-027-632-240569	Sequence 240569, S
41	99.2	2.7	650	14	US-10-027-632-240569	Sequence 240569, S
42	99.2	2.7	98	13	US-10-027-633-122186	Sequence 122186, S
43	99.2	2.7	98	14	US-10-027-632-122186	Sequence 122186, S
44	99.2	2.7	1260	13	US-10-027-633-240568	Sequence 240568, S
45	99.2	2.7	1260	14	US-10-027-632-240568	Sequence 240568, S

ALIGNMENTS

```

RESULT 1
US-09-900-448-3
; Sequence 3, Application US/09900448
; Publication No. US20030220488A1
; GENERAL INFORMATION:
; APPLICANT: CECCARDI, Toni et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001272
; CURRENT APPLICATION NUMBER: US/09/900,448
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13737
; TYPE: DNA
; ORGANISM: Human
; US-09-900-448-3

```

Query Match	Score	DB	Length
100.0%	3737	13	13737

Matches 3737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTCAAGAGAGAGGTCAAGTCAAGCTACCAAGAACAGATTGGAAAGTCATCAGCATTTAGT	60
Db	10001	CTCAAGAGAGAGGTCAAGTCAAGCTACCAAGAACAGATTGGAAAGTCATCAGCATTTAGT	10060
Qy	61	GGCATTTTAAAGCCCTCGAGCTAGGTGAGATTACCAAGGAAGTGAAGGTGAGAGAGA	120
Db	10061	GGCATTTTAAAGCCCTCGAGCTAGGTGAGATTACCAAGGAAGTGAAGGTGAGAGAGAA	10120
Qy	121	GAAGAGGCCCAAGTAGAGGAGATTCCAATTTTAGATATCAGGTGGAAGAAAAAGTAGTC	180
Db	10121	GAAGAGGCCCAAGTAGAGGAGATTCCAATTTTAGATATCAGGTGGAAGAAAAAGTAGTC	10180
Qy	181	AAAAAAGATGAGGAATTCTGGAGAGTCAAGGTGCACAGAAGCCAAGTTCAAAAAAA	240

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 197
LENGTH: 1282
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (675)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1195)
OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-197

Query Match 1.9%; Score 69.2; DB 4; Length 1282;
Best Local Similarity 73.0%; Pred. No. 1.6e-11;
Matches 89; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 3554 GCTTGTTCAGTCTAGCTGCTATTAAGAAATATCTTAGAGTGGTAATCTATCA 3613
DB 505 GGTATGGTATCTTCTTGTGCTATTAACAAATACAGAGACTGGTAAATTGTAA 446
QY 3614 GCATAGCAATTTATTTTCAATCTGAGGCTGGAATCAAGATCAAGGCTCAG 3673
DB 445 GGAACAGAAATGTAATTTCTACACATCTTGAGAGATGGAGAGATCAAGAGCTGG 386
QY 3674 CA 3675
DB 385 CA 384

Search completed: December 16, 2003, 05:46:27
Job time: 161.331 secs

[illegible]

```

1  PRIOR APPLICATION NUMBER: 60/092122
2  PRIOR FILING DATE: 1998-07-09
3
4  Query Match 1.9%; Score 69.8; DB 4; Length 762;
5  Best Local Similarity 71.3%; Pred. No. 7e-12;
6  Matches 92; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
7
8  QY 3609 TATCAGCAATGAAATTTATTGTTCCACATTTCTGAGAGGCTGGAAATTCACAGTCAAGGC 3668
9  DB 756 TATTAACAGCAAAATATTATTATTCTTACAGTTCTAGAGGCTGGAGATCCAAATCAAGGC 697
10 QY 3669 TCCAGCAGGTTCTAGTCTGCTGAGGCTTGTCTGCTTCGAAGATGAGACCTTTTNGCT 3728
11 DB 696 ACCAGCAGATTCAGATCTGTGTAAGACCAAGTTCTTCTCTGACAGTGCCTTAGCT 637
12 QY 3729 GTGTTCTCA 3737
13 DB 636 GTGATCACA 628
14
15 RESULT 15
16 US-09-205-258-197/c
17 Sequence 197, Application US/09205258
18 Patent No. 6525174
19
20 -GENERAL INFORMATION:
21 APPLICANT: Young et al.
22 TITLE OF INVENTION: 207 Human Secreted Proteins
23 FILE REFERENCE: P2007P1
24 CURRENT APPLICATION NUMBER: US/09/205,258
25 CURRENT FILING DATE: 1998-12-04
26 EARLIER APPLICATION NUMBER: PCT/US98/11422
27 EARLIER FILING DATE: 1998-06-04
28 EARLIER APPLICATION NUMBER: 60/048,885
29 EARLIER FILING DATE: 1997-06-06
30 EARLIER APPLICATION NUMBER: 60/049,375
31 EARLIER FILING DATE: 1997-06-06
32 EARLIER APPLICATION NUMBER: 60/048,881
33 EARLIER FILING DATE: 1997-06-06
34 EARLIER APPLICATION NUMBER: 60/048,880
35 EARLIER FILING DATE: 1997-06-06
36 EARLIER APPLICATION NUMBER: 60/048,896
37 EARLIER FILING DATE: 1997-06-06
38 EARLIER APPLICATION NUMBER: 60/049,020
39 EARLIER FILING DATE: 1997-06-06
40 EARLIER APPLICATION NUMBER: 60/048,876
41 EARLIER FILING DATE: 1997-06-06
42 EARLIER APPLICATION NUMBER: 60/048,895
43 EARLIER FILING DATE: 1997-06-06
44 EARLIER APPLICATION NUMBER: 60/048,884
45 EARLIER FILING DATE: 1997-06-06
46 EARLIER APPLICATION NUMBER: 60/048,894
47 EARLIER FILING DATE: 1997-06-06
48 EARLIER APPLICATION NUMBER: 60/048,971
49 EARLIER FILING DATE: 1997-06-06
50 EARLIER APPLICATION NUMBER: 60/048,964
51 EARLIER FILING DATE: 1997-06-06
52 EARLIER APPLICATION NUMBER: 60/048,882
53 EARLIER FILING DATE: 1997-06-06
54 EARLIER APPLICATION NUMBER: 60/048,899
55 EARLIER FILING DATE: 1997-06-06
56 EARLIER APPLICATION NUMBER: 60/048,893
57 EARLIER FILING DATE: 1997-06-06
58 EARLIER APPLICATION NUMBER: 60/048,900
59 EARLIER FILING DATE: 1997-06-06
60 EARLIER APPLICATION NUMBER: 60/048,901
61 EARLIER FILING DATE: 1997-06-06
62 EARLIER APPLICATION NUMBER: 60/048,892
63 EARLIER FILING DATE: 1997-06-06
64 EARLIER APPLICATION NUMBER: 60/048,915
65 EARLIER FILING DATE: 1997-06-06
66 EARLIER APPLICATION NUMBER: 60/049,019
67 EARLIER FILING DATE: 1997-06-06
68 EARLIER APPLICATION NUMBER: 60/048,970

```

Db 1184 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1243
Qy 744 TTTCAGCAGGAGGACAGTGCACCATTCGTCGACCAAGTGTCTATTCTCTCTC 803
Db 1244 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1303
Qy 804 ACCTGGCCTTCCTCATCTGCGCTGAGATGATCTCTCCCTCCTCAGCACTATCTG 863
Db 1304 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1363
Qy 864 CATTCATCATACCTCTTCTCTGCGCTGAGATGATCTCTCCCTCCTCAGCACTATG 923
Db 1364 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1423
Qy 924 GATTCACAGGAGGACCATTCGCTGAGACACGCGGAGATGCTG 975
Db 1424 YYYYYYYYYYGTACCAATTCCTATCTTAACTTAACTGATGATG 1475

RESULT 14
US-09-996-243-344/C
Sequence 344, Application US/09996243
Patent No. 6478825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Batton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: KJawin, Ivar J.
APPLICANT: Nadjler, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C13
CURRENT APPLICATION NUMBER: US/09/996,243
PRIOR APPLICATION NUMBER: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801

Query Match 2.1%; Score 78.4; DB 3; Length 8353;
Best Local Similarity 66.8%; Pred. No. 6.5e-14;
Matches 127; Conservative 0; Mismatches 61; Indels 2; Gaps 1;

QY 3547 CAGAAATGGTTGTTTCAGTCAGTTAGGCTGCTATAGAAATATCTTAGAGTGGTAA 3606
DB 8244 CAGTATTTGTTATTTGTTAGTCGAGATGTTATTAACAAATACAGCACCTGGGTGG 8185
QY 3607 TCTATCAGCAATAGCAATTTATTTGTTACAAATTCGAGGCTGGAATTCAGATTAAG 3666
DB 8184 CTTATTAACAAATAGAAATGTTCTTAGAGTCCTGAGGAGTGGAAATCCAATCAAT 8125
QY 3667 GCTCCAGCAGGTCAGTGTGCTGAGTGTGTT--CTGCTCGAAGTGGCACCTTTT 3724
DB 8124 GCACCAACAGCTTGTGTGTCGTGAGGGCCAGTTCTCTGTTCAATGATGATGCTTGT 8065
QY 3725 TGCTGTGTTTC 3734
DB 8064 CACTGGGCTC 8055

RESULT 12
US-09-009-913-1/c
; Sequence 1, Application US/0909913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: Axy's Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-Jan-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3231
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72928 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-09-009-913-1

Query Match 2.1%; Score 77.6; DB 3; Length 72928;
Best Local Similarity 72.2%; Pred. No. 5.6e-13;
Matches 130; Conservative 0; Mismatches 44; Indels 6; Gaps 2;
QY 3560 TTTCAGTCAGTTAGGCTGCTATAGCAATATCTTAGAGTGGTATCTATAGCAATA 3619
DB 3623 TCTTATGTCATTAACTGACCAAAAAGATGCTATAGCTAGG----CTACAGCAAGC 36208

QY 3620 GGAATTTATGTTTACAAATTCGAGCTGGAATATCCAGATCAAGGCTCCAGAGTT 3679
DB 36207 GAAATTTATTTATTCAGAAATTTCTAGAGCTGGAGATCCAGACCAAGCAGATT 36148
QY 3680 CAGTGTCTGCTAGTGS--CTTGTCTGCTTGAAGATGCGACCTTTTGTGTCTCA 3737
DB 36147 CAGCGTGTGTGAGGGCCACCTTCTTAATGATGATGATCTTCTCACCGTACCTCA 36088

RESULT 13
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZ9PT-F18
; US-08-232-463-14

Query Match 1.9%; Score 70; DB 1; Length 7218;
Best Local Similarity 5.3%; Pred. No. 3e-11;
Matches 22; Conservative 235; Mismatches 155; Indels 0; Gaps 0;
QY 564 GAAGTGTCCTCTTTTAAGTTCGACACTGATTTCTATTAATCAATTAACCGATCTC 623
DB 1064 GATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1123
QY 624 TGTACATGACACCTGAATCTTAGAGAGTGCAGCCCAACCCAGATGTTGCGCTTAC 683
DB 1124 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1183
QY 684 CTGGAATCTAGCACTGTTTTCACACTGCTTCTTTCAGGCACTGCTGATTCAG 743

OTHER INFORMATION: 12-500-220 : polymorphic base A or G
NAME/KEY: misc binding
LOCATION: 481..500
OTHER INFORMATION: 12-500-220.mis1, potential
NAME/KEY: misc binding
LOCATION: 502..521
OTHER INFORMATION: 12-500-220.mis2, potential complement
NAME/KEY: primer bind
LOCATION: 283..303
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer bind
LOCATION: 711..731
OTHER INFORMATION: downstream amplification primer, complement
NAME/KEY: misc binding
LOCATION: 489..513
OTHER INFORMATION: 12-500-220 potential probe
US-09-671-317-212

Query Match 2.1%; Score 79; DB 4; Length 1001;
Best Local Similarity 75.5%; Pred. No. 9.1e-15;
Matches 11; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

Qy 3558 TGTTCAGTCAGTTAGGCTGCTATAGA-GAATATCTTAGGCTTAATCATCAGCA 3616
Db 837 TGTCTGGTTTGTGGGCTGCTGTAACAAATACCTTACAGTGGCTTAATAACA 896
Qy 3617 ATAGAAATTATTTGTCACAAATTCGAGGCTGGAATCAAGATCAAGCTCCAGCAG 3676
Db 897 ACAGAAATGATTTGCTCAGCTTCAGAGCTGGGAAGTTCAAGATCAAGCTCCAGCAG 956
Qy 3677 GTTCAGTGTCTGCTGAGTGTCTTCT 3703
Db 957 TTTCGGTGTGTGTGAGAGCTTTTGTCT 983

RESULT 10
US-09-671-317-458
Sequence 458, Application US/09671317
Patent No. 6528260
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marra
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLERIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
FILE REFERENCE: 62.US3.CIP
CURRENT APPLICATION NUMBER: US/09/671.317
CURRENT FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 458
LENGTH: 1001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 501
OTHER INFORMATION: 12-500-217 : insertion CATATA
NAME/KEY: misc binding
LOCATION: 481..500
OTHER INFORMATION: 12-500-217.mis1, potential
NAME/KEY: primer bind
LOCATION: 286..306
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer bind

LOCATION: 714..734
OTHER INFORMATION: downstream amplification primer, complement
US-09-671-317-458

Query Match 2.1%; Score 79; DB 4; Length 1001;
Best Local Similarity 75.5%; Pred. No. 9.1e-15;
Matches 11; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

Qy 3558 TGTTCAGTCAGTTAGGCTGCTATAGA-GAATATCTTAGGCTTAATCATCAGCA 3616
Db 840 TGTCTGGTTTGTGGGCTGCTGTAACAAATACCTTACAGTGGCTTAATAACA 899
Qy 3617 ATAGAAATTATTTGTCACAAATTCGAGGCTGGAATCAAGATCAAGCTCCAGCAG 3676
Db 900 ACAGAAATGATTTGCTCAGCTTCAGAGCTGGGAAGTTCAAGATCAAGCTCCAGCAG 959
Qy 3677 GTTCAGTGTCTGCTGAGTGTCTTCT 3703
Db 960 TTTCGGTGTGTGTGAGAGCTTTTGTCT 986

RESULT 11
US-08-611-587-1/c
Sequence 1, Application US/08611587
Patent No. 6150091
GENERAL INFORMATION:
APPLICANT: PANDOLFO, MASSIMO
APPLICANT: MONTERMINI, LAURA
APPLICANT: MOLTO, MARIA D.
APPLICANT: Koenig, Michael
APPLICANT: Campuzano, Victoria
APPLICANT: Cossee, Mireille
TITLE OF INVENTION: Direct Diagnosis of Friedreich Ataxia
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: Fulbright & Jaworski L.L.P. Patent Dept.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.
ZIP: 77010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,587
FILING DATE: 03-MAR-1996
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Blashear-Macatee, Sarah J.
REGISTRATION NUMBER: 38,087
REFERENCE/DOCKET NUMBER: D-5901
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5620
TELEFAX: 713-651-5246
TELEX: 76-2829
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 9q13
UNITS: bp
US-08-611-587-1

SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 174493
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)....(174493)
OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3

Query Match 2.2%; Score 82.6; DB 4; Length 174493;
Best Local Similarity 72.2%; Pred. No. 2.6e-14;
Matches 135; Conservative 0; Mismatches 49; Indels 3; Gaps 2;

QY 3551 AATGGTTGTTGACGAGTTAGGCTGCTATAGAGAAATCTTAAGGCGGTAATCTA 3610
DB 14097 AAAGTGCATCTGAGTCTGTTGGCTGCGCATATAG-ATACCTTAGACTTGGCAATTTA 14155
QY 3611 TCAGCAATAGGAATTTATTTGTTCCAAATCTGAGGCTGGAATAATCAAGTCAAGGCTC 3670
DB 14156 TAAACATATAGAAATTTCTTCTGACAGTTGTGAAGACTGGGAAGTCCAAAGTCAAGGCGC 14215
QY 3671 CAGCAGGTTGAGTGTCTGCTGAGTGTCTGCTTCTGCTTGAAGATGGACCTTTTCTGT 3730
DB 14216 CAGCAGATCTGTATCTGTGATGAGTCT--CCCTGCTTCAAAAATGGCGCTTCTGTCTGC 14273
QY 3731 GTTCTCA 3737
DB 14274 ATCTTCA 14280

RESULT 7

US-09-873-404-3
Sequence 3, Application US/09873404
Patent No. 6500656
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001212-CIP
CURRENT APPLICATION NUMBER: US/09/873,404
CURRENT FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 63588
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)....(63588)
OTHER INFORMATION: n = A,T,C or G
US-09-873-404-3

Query Match 2.2%; Score 80.8; DB 4; Length 63588;
Best Local Similarity 71.6%; Pred. No. 4.7e-14;
Matches 106; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 3550 AATGGTTGTTGACGAGTTAGGCTGCTATAGAGAAATCTTAAGGCGGTAATCT 3609
DB 15658 AATCATCTGTTTATGCTATTTCCAGAGCTATTAACAAATACATTAAGTGGCTT 15717
QY 3610 ATCAGCAATAGGAATTTATTTGTTCCAAATCTGAGGCTGGAATAATCAAGTCAAGGCT 3669
DB 15718 ATAAACAGCAGAAATTTATTTCTGSCAGTCTGGAAGCTGGAAAGTCAAGATCAAGGTG 15777
QY 3670 CAGCAGGTTGAGTGTCTGCTGAGTGTCT 3697
DB 15778 CAGGCAATTCAGTGTCTGCTGAGGTT 15805

RESULT 8

US-09-146-053-4
Sequence 4, Application US/09146053A
Patent No. 6399349
GENERAL INFORMATION:
APPLICANT: Ryan, James W.
APPLICANT: Sprinkle, Terry Joe Curtis
APPLICANT: Venema, Richard C.
TITLE OF INVENTION: Human Aminopeptidase P Gene
FILE REFERENCE: MCG103
CURRENT APPLICATION NUMBER: US/09/146,053A
CURRENT FILING DATE: 1998-09-02
EARLIER APPLICATION NUMBER: 60/057,854
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 50000
TYPE: DNA
ORGANISM: Homo sapiens
US-09-146-053-4

Query Match 2.2%; Score 80.4; DB 4; Length 50000;
Best Local Similarity 68.7%; Pred. No. 5.4e-14;
Matches 125; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY 3556 TTGTTTCACTGAGTTAGGCTGCTATAGAGAAATCTTAAGGCGGTAATCTATCAGC 3615
DB 37896 TTATCTTATGATCTAGTGGGCTGCTATTAACAAAGGCTGCAAGCTGTTGCTTAATTAAC 37955
QY 3616 AATAGAAATTTATTTGTTCCAAATCTGAGGCTGGAATAATCAAGTCAAGGCTCCAGCA 3675
DB 37956 AGTAGACATTTATTTCTTACAGTCTGAGGCTGGGAATTCAGAAATCAAGGACTAGCA 38015
QY 3676 GGTTCAGTGTCTGCTGAGTGTCTTCTGCTTGAAGATGGACCTTTTCTGTCTGTCT 3735
DB 38016 GATTGTGTCTCCAAAGGCGCCAGT-TCCTTGTATGATGGCACCTCTAGCTGTATCTCT 38074
QY 3736 CA 3737
DB 38075 CA 38076

RESULT 9

US-09-671-317-212
Sequence 212, Application US/09671317
Patent No. 6528260
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bouguetelret, Lydie
APPLICANT: Cohen, Amick
TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
FILE REFERENCE: 62,US3,CIP
CURRENT APPLICATION NUMBER: US/09/671,317
CURRENT FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 212
LENGTH: 1001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 501

OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29981
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30136
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30140
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31205
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31206
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31592
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33095
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33160
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34066
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34072
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 36816
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 39020
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42164
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42459
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46808
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46823
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46826
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 47291
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52786
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52787
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 53384
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 54684
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59215
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59235
OTHER INFORMATION: unknown

NAME/KEY: unsure
LOCATION: 59242
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 63290
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 66614
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68660
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68697
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68718
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68733
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68739
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 69785
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79134
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79198
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 86336
OTHER INFORMATION: unknown
OTHER INFORMATION: unknown

US-09-791-211-3

Query Match 2.3%; Score 86; DB 4; Length 87543;
Best Local Similarity 73.3%; Pred. No. 1.2e-15;
Matches 151; Conservative 0; Mismatches 50; Indels 5; Gaps 3;

QY 3534 TCCAGGAACTCCCAAAATGTTTTCAGTCACTTAGCTGCT---ATAGAGAAAT 3590
DB 68969 TACTAGTAGTGCAGACACTGATTGTCTAGTCTGTTTGTCTTCATACAAAT 68910
QY 3591 AFTTAGTGGGTATCTATCAGCAATTAATTTTTCACATTTCTGAGGCTGG 3650
DB 68909 GCCACAGACTGGGCAATTTTCAACACACAGTAGTTTATTTCTCATAGTTCTGAGGCTGG 68850
QY 3651 AAATCCAAATCAAGGCTCCAGCAGTTCACTGCTCTGCTGAGTCTTGT-TCGCTTTCG 3709
DB 68849 -AGGTCCAGATCAAGGCGCCAGTAGTTCACTGCTCTGCTGAGGCTGCTCTTCC 68791
QY 3710 AAGATGCACCTTTTGTCTGCTCT 3735
DB 68790 AAATGAGGCTTGTCTGCTGCTCT 68765

RESULT 6
US-09-804-471A-3
Sequence 3, Application US/09804471A
Patent No. 6479269

GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL0001164
CURRENT APPLICATION NUMBER: US/09/804,471A
CURRENT FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 4

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 87350 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-79

Query Match 2.3%; Score 86; DB 3; Length 87350;
Best Local Similarity 73.3%; Pred. No. 1.2e-15;
Matches 151; Conservative 0; Mismatches 50; Indels 5; Gaps 3;

QY 3534 TCCAGGGAATCCCAAAATGTTTTCAGTCAGTTAGGCTGCT--ATAAGAAAT 3590
DB 68776 TACTAGTATGTCAGACAGATGATTTGCTAGTCTGTTTGTCTTCAATACAAAT 68717
QY 3591 ATCTTAGAGTGGTAATCTATTCAGCAATAGAAATTTATGTTCAATTCGAGGCTGG 3650
DB 68716 GCCACAGACTGGGCAATTTACAACAACAGATGTTATTTCTCATAGTTCTGAGGCTGG 68657
QY 3651 AATATCAAGATCAAGGCTCCAGAGGTTCAAGTGTCTGCTGAGTCTGTG-TCTGCTTCG 3709
DB 68656 -AGTCCAGATCAAGGCGCAGTAGGTTCAAGTGTCTGAGGCTGCTCTGCTTCC 68598
QY 3710 AAGATGACACCTTTTGTCTGTTCT 3735
DB 68597 AATATGAGGCTTGTCTGCTGTCT 68572

RESULT 4

US-09-618-166-79/C
Sequence 79, Application US/09618166
Patent No. 6583112
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
Oshima, Junko
Mulligan, John T.
Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 87350 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 79:
US-09-618-166-79

Query Match 2.3%; Score 86; DB 4; Length 87350;
Best Local Similarity 73.3%; Pred. No. 1.2e-15;
Matches 151; Conservative 0; Mismatches 50; Indels 5; Gaps 3;

QY 3534 TCCAGGGAATCCCAAAATGTTTTCAGTCAGTTAGGCTGCT--ATAAGAAAT 3590
DB 68776 TACTAGTATGTCAGACAGATGATTTGCTAGTCTGTTTGTCTTCAATACAAAT 68717
QY 3591 ATCTTAGAGTGGTAATCTATTCAGCAATAGAAATTTATGTTCAATTCGAGGCTGG 3650
DB 68716 GCCACAGACTGGGCAATTTACAACAACAGATGTTATTTCTCATAGTTCTGAGGCTGG 68657
QY 3651 AATATCAAGATCAAGGCTCCAGAGGTTCAAGTGTCTGCTGAGTCTGTG-TCTGCTTCG 3709
DB 68656 -AGTCCAGATCAAGGCGCAGTAGGTTCAAGTGTCTGAGGCTGCTCTGCTTCC 68598
QY 3710 AAGATGACACCTTTTGTCTGTTCT 3735
DB 68597 AATATGAGGCTTGTCTGCTGTCT 68572

RESULT 5

US-09-791-211-3/C
Sequence 3, Application US/09791211
Patent No. 6448080
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
FILE REFERENCE: RTS-0205
CURRENT APPLICATION NUMBER: US/09/791,211
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 3
LENGTH: 87543
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 7421
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 7427
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 11609
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12605
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12742
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 23570
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29422
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29979
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29980

```

RESULT 2
US-09-539-333D-177
; Sequence 177, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueterec, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essiloux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSER.0472AUS
CURRENT APPLICATION NUMBER: US/09/539,333D
CURRENT FILING DATE: 2000-03-30
PRIORITY APPLICATION NUMBER: US 60/126,903
PRIORITY FILING DATE: 1999-03-30
PRIORITY APPLICATION NUMBER: US 60/131,971
PRIORITY FILING DATE: 1999-04-30
PRIORITY APPLICATION NUMBER: US 60/132,065
PRIORITY FILING DATE: 1999-04-30
PRIORITY APPLICATION NUMBER: US 60/143,928
PRIORITY FILING DATE: 1999-07-14
PRIORITY APPLICATION NUMBER: US 60/145,915
PRIORITY FILING DATE: 1999-07-27
PRIORITY APPLICATION NUMBER: US 60/146,453
PRIORITY FILING DATE: 1999-07-29
PRIORITY APPLICATION NUMBER: US 60/146,452
PRIORITY FILING DATE: 1999-07-29
PRIORITY APPLICATION NUMBER: US 60/162,288
PRIORITY FILING DATE: 1999-10-28
PRIORITY APPLICATION NUMBER: US 09/416,384
PRIORITY FILING DATE: 1999-10-12
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patent.pm
SEQ ID NO 177
;
; LENGTH: 3001

```

```

Query Match      2.4%; Score 88; DB 4; Length 3001;
Best Local Similarity 71.7%; Pred. No. 2.5e-17;
Matches 129; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

QY      3559 GTTTCAGTCAGTTTAGGCTGCTATAGAGAAATCTTAGAGTGGGTATCTATCAGCAAT 3618
DB      2055 GTTCTGATCGATTTTGGTGGCTGCTGTAACAGATACCCAGACCTGGGTATTAATTAAGAAAG 2114
QY      3619 AGGAATTTATTTGTTCCAAATTTCTGAGAGCTGGAATAATCCAAAGATCAAGGCTCCAGCAGGT 3678
DB      2115 AAGAAATTCATTTCTCATAGTTGTGAGAGCTGAGAAAGTCCAGATGAAAGGCCCAAGCAAT 2174
QY      3679 TCAAGTGTCTGCTGAGTGCT-TGTTCTGCTTCGAAGATGGCACTTTTGTCTGTGTTCTCA 3737
DB      2175 TTGGTGTGTGGTAGGGGTTCCGCTCTGCTTCGAAGATGAGCGCTTGATGTGCAATTTTCA 2234

RESULT 3
US-08-781-891-79/c
Sequence 79, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-Bn
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 11:21:38 ; Search time 158.331 Seconds
(Without alignments)
10417.750 Million cell updates/sec

Title: US-09-900-448-3_COPY_10001_13737

Perfect score: 3737
Sequence: 1 cccaaggagagcagtcacga.....accttttgctgtgtctca 3737

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued Patents NA:*
2: /cgn2_6/prodata/2/ina/5A COMB.seq:*
3: /cgn2_6/prodata/2/ina/5B COMB.seq:*
4: /cgn2_6/prodata/2/ina/6A COMB.seq:*
5: /cgn2_6/prodata/2/ina/6B COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	2.5	7218	1	US-08-232-463-14
2	88	2.4	3001	4	US-09-539-333D-177
3	86	2.3	87350	3	US-08-781-891-79
4	86	2.3	87350	4	US-09-618-166-79
5	86	2.3	87543	4	US-09-791-211-3
6	82.6	2.2	174493	4	US-09-804-471A-3
7	80.8	2.2	63588	4	US-09-873-404-3
8	80.4	2.2	50000	4	US-09-146-053-4
9	79	2.1	1001	4	US-09-671-317-212
10	79	2.1	1001	4	US-09-671-317-458
11	78.4	2.1	8353	3	US-08-611-587-1
12	77.6	2.1	72928	3	US-09-009-913-1
13	70	1.9	7218	1	US-08-232-463-14
14	69.8	1.9	762	4	US-09-596-243-344
15	69.2	1.9	1282	4	US-09-205-258-197
16	69	1.8	50000	4	US-09-146-053-3
17	68.4	1.8	197486	4	US-09-877-177A-10
18	68.2	1.8	309	4	US-09-222-575-145
19	68.2	1.8	309	4	US-09-389-681-145
20	68.2	1.8	309	4	US-09-620-405B-145
21	68.2	1.8	309	4	US-09-339-338-145
22	68.2	1.8	309	4	US-09-433-826B-145
23	68.2	1.8	309	4	US-09-604-287A-145
24	67.8	1.8	50000	4	US-09-146-053-4
25	66.8	1.8	41684	4	US-09-536-059-1
26	66.6	1.8	1001	4	US-09-641-638-459
27	63.2	1.7	246240	2	US-08-724-394A-20

28	63.2	1.7	246240	2	US-08-724-394A-21	Sequence 21, Appl
29	63.2	1.7	246240	2	US-08-724-394A-22	Sequence 22, Appl
30	63	1.7	3001	4	US-09-539-333D-192	Sequence 192, App
31	62.8	1.7	44453	4	US-09-146-053-5	Sequence 5, Appl1
32	62	1.7	16857	4	US-09-426-290-1	Sequence 1, Appl1
33	62	1.7	319608	4	US-09-539-333D-1	Sequence 1, Appl1
34	62	1.7	319608	4	US-09-679-409-1	Sequence 1, Appl1
35	61.8	1.7	152331	3	US-09-128-155-16	Sequence 16, Appl
36	61.8	1.7	176373	3	US-09-128-155-17	Sequence 17, Appl
37	61.4	1.6	8355	3	US-08-406-030A-23	Sequence 23, Appl
38	60.4	1.6	3001	4	US-09-539-333D-138	Sequence 138, App
39	59.4	1.6	202001	4	US-09-734-674-3	Sequence 3, Appl1
40	59	1.6	80246	3	US-09-078-294-4	Sequence 4, Appl1
41	59	1.6	80595	3	US-09-078-294-3	Sequence 3, Appl1
42	57	1.5	49312	4	US-09-671-317-485	Sequence 485, App
43	55	1.5	116592	4	US-09-818-512-3	Sequence 3, Appl1
44	53.6	1.4	49136	3	US-09-422-869-1	Sequence 1, Appl1
45	52	1.4	43069	4	US-09-292-542A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZPc-F1s
US-08-232-463-14
Query Match 2.5%; Score 95; DB 1; Length 7218;

THIS PAGE BLANK (USPTO)

CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.

CC Sequence 228 BP; 53 A; 58 C; 60 G; 57 T; 0 other;

QY Query Match 3.0%; Score 111.2; DB 25; Length 228;
 Db Best Local Similarity 79.9%; Pred. No. 1.3e-23;
 Matches 131; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1158 GATATGCTTCTCTGACAAAGGAGGCTATACCTAGTAGCGGTTATCCGAAGCGGCTG 1217
 Db 1 GTATATATCTTCTCTGACAAAGGAGGCTATACCTAGTAGCGGTTATCCGAAGCGGCTG 60

QY 1218 GAGAGGAAAGTCCGGAACCCCTCATGGGATTATCTGTGACTGTGTGAGCGGCTTTATC 1277
 Db 61 GAGAGGAAATTTGGAGAGCCCTGATGGGGTCTGCTTCATTCTGTGATGAGGCTTTACC 120

QY 1278 TGCCCTGGGCTTCTCTGCTCCATATCATATGCGAGGTGAGGCGCT 1321
 Db 121 TGCTCTGATCTTCTCTGAGCTCTACATCATGCGAGGCGCAAGCT 164

RESULT 15

ID ABX47771 standard; cDNA; 397 BP.

AC ABX47771;

DT 21-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #12936.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;

KM muscle deposition; fat deposition; genome mapping; gene identification;

KM Gene analysis; cattle breeding.

OS Bos Taurus.

PN US2002137139-A1.

PD 26-SEP-2002.

PF 24-SEP-2001; 2001US-0960352.

PR 12-JAN-1999; 99US-115707P.

PA 11-JAN-2000; 2000US-0480902.

PA (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

PI Byatt JC, Mathialagan N, Tao N, Warren WC;

DR MPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat
 XX deposition, useful for genome mapping, gene identification and
 XX analysis, cattle breeding, or for genetically improving cattle
 XX Claim 2; SEQ ID No 12936; 245bp; English.
 CC The invention relates to a purified nucleic acid molecule associated with

CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are; (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.

CC Sequence 397 BP; 85 A; 115 C; 99 G; 98 T; 0 other;

QY Query Match 2.8%; Score 106; DB 25; Length 397;
 Db Best Local Similarity 73.1%; Pred. No. 7.9e-22;
 Matches 136; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1136 CTTTGCTCTCCAGGAGCACCAGTATATGCTTCTCTGACAAAGGAGGCTATACCTAGT 1195
 Db 211 CTACCTCAGCCAGCGGACGATATATATGCTCTGACAAAGGAGGCTATACCTAGT 270

QY 1196 AAGCGTTATCCGAAGCGGCTGGAAGAAGCGGAGCCCTCATGGGATTATCTGGA 1255
 Db 271 GCAGATTATCCAGCAAGCAAGCTGGAAGACCATTTGGAGCCCTCATGGGATTATCTGGA 330

QY 1256 CTCTGTGATGCGGCGCTTTATCTGCGCTGGGCTTCTCGGCTCATATCATGCGAGTGA 1315
 Db 331 TTCTGTGGCTGAGGCTTTATCTGCGCTGGGCTTCTCGGCTCATATCATGCGAGTGA 390

QY 1316 GGGGCT 1321
 Db 391 GAAAGCT 396

Search completed: December 15, 2003, 12:36:53
 Job time : 592.508 secs

KW gene analysis; cattle breeding.
 XX Bos Taurus.
 OS US2002137139-A1.
 PN 26-SEP-2002.
 PD 24-SEP-2001; 2001US-0960352.
 XX 12-JAN-1999; 99US-115707P.
 PR 11-JAN-2000; 2000US-0480902.
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 PI WPI; 2003-110599/10.
 DR New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX
 PS Claim 2; SEQ ID No 6846; 245bp; English.
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are: (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.
 CC
 XX
 SO Sequence 383 BP; 85 A; 108 C; 94 G; 96 T; 0 other;
 Query Match 3.1%; Score 117.4; DB 25; Length 383;
 Best Local Similarity 63.7%; Pred. No. 2.2e-25;
 Matches 244; Conservative 0; Mismatches 56; Indels 83; Gaps 1;
 QY 927 TTCCACAGGAGCCACTACTGCGGTGACACACAGCGGAGATGGCGATAGCTGGCC 986
 DB 84 TTTCAGTGAACACCTACTGCGGTGACACACAGCGGAGATGGCGATAGCTGGCC 143
 QY 987 CATTGCTCATCTGCGGCCAGGCTCTTTCAGACAGTATGCTCTTTCTCTGGAGAA 1046
 DB 144 CATTGAGCATCTGCGGCCAGGCTCTTTCAGACAGTATGCTCTTTCTCTGGAGTA 203
 QY 1047 AAAACCTCTATCTGCTCAAGGTGTGTATTTGGGGAGAGGCTTGAAGTGAAGCTGGACAA 1106
 DB 204 AAGCTCTTAC----- 214

QY 1107 GCATATCAACTCTGTATTATTATACCATCTTGTCTCCAGGGACCCAGATATATGTC 1166
 DB 215 -----TATCCAGGCTACCCAGGATATATTC 240
 QY 1167 TTCTGACAAAGGAGGCTATACCTTAGTAAGCGGTTATCCGAGCGCTGAGAAAGAA 1226
 DB 241 TTCTGACAAAGGAGGCTATACCTTAGTAAGATTTATCCATAGCAGCTGAGAAAGAA 300
 QY 1227 GTGCGAGCCCTCATGAGATATCTGACTCTGTGATGCGGCTTTATCTCCCTGGG 1286
 DB 301 TTGCGAGCCCTCATGAGGCTGTGCTTCATTTGTGATGACAGCTTATCTGTCTGGA 360
 QY 1287 TCTTCTGCGCTCATATCATGCGC 1309
 DB 361 TCTTCTGCGCTCATATCATGCGC 383
 RESULT 14
 ABX43075
 ID ABX43075 standard; cDNA; 228 BP.
 AC ABX43075;
 XX 20-FEB-2003 (first entry)
 DT
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #8240.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX US2002137139-A1.
 XX 26-SEP-2002.
 PD 24-SEP-2001; 2001US-0960352.
 XX 12-JAN-1999; 99US-115707P.
 PR 11-JAN-2000; 2000US-0480902.
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 PI WPI; 2003-110599/10.
 DR New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX
 PS Claim 2; SEQ ID No 8240; 245bp; English.
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are: (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the

CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMPD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139.
CC
XX
SQ Sequence 420 BP; 90 A; 114 C; 116 G; 100 T; 0 other;

Query Match 3.3%; Score 123.6; DB 25; Length 420;
Best Local Similarity 79.0%; Pred. No. 2.8e-27;
Matches 147; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1136 CTTTGTCTCCAGGACCCAGGATATGCTTCTCGACCAAGGAGGCTTATCCCTAGT 1195
DB 93 CTACCTCATCAGGGTACCAAGATATATCTTCTCGACCAAGGAGGCTTATCCCTAGT 152
QY 1196 AACGGCTTATCCGAGCGCTGAGAGAGAGTCCGAGCCCTCATGGAATTATCTGGA 1255
DB 153 AAAAGATTATCCAAAGAGCTGAGAGAGATTTGGAGGCTCATGAGGCTTCTGCTTCA 212
QY 1256 CTTCTGTGATGCGGCTTATCTGCTGCGCTCTTCTCGCTCCATATCATGAGAGTGA 1315
DB 213 TTTCTGTGAGTACGAGCTTATCTGCTGATCTTCTCACTCATCATGAGAGCTCA 272
QY 1316 GGGGCT 1321
DB 273 GAAGCT 278

RESULT 12
ABX40203
ID ABX40203 standard; cDNA; 374 BP.
AC ABX40203;
XX
XX 20-FEB-2003 (first entry)
DT
XX
DE Bovine EST associated with lactation/muscle/fat deposition #5368.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
OS
XX Bos Taurus.
XX
XX US2002137139-A1.
XX
XX 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-0960352.
XX
XX 12-JAN-1999; 99US-115707P.
XX 11-JAN-2000; 2000US-0480902.
XX
XX (BYAT/) BYAT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARRE/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
XX WPI, 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and
XX analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2; SEQ ID No 5368; 245bp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMPD), derived

CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49847, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMPD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMPD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139.
CC
XX
SQ Sequence 374 BP; 86 A; 102 C; 92 G; 94 T; 0 other;

Query Match 3.2%; Score 118.2; DB 25; Length 374;
Best Local Similarity 64.1%; Pred. No. 1.2e-25;
Matches 243; Conservative 0; Mismatches 53; Indels 83; Gaps 1;

QY 927 TTCCACAGGAGACCACTACCTGCGCTGACACACAGCCGGAGTGGCTGATGCTGGCC 986
DB 79 TTTCAGTGAAGACCACTACCTGCGCTGACACACAGCCGGAGTGGCTGATGCTGGCC 138
QY 987 CATTCCTCATAGTGGAGCCCGAGGCTCTTACAGAGTGAATGCTCTTTCTCGGAGAA 1046
DB 139 CATTGAGCATCTGTGAGCCCGAGGCTCTTACAGAGTGAATGCTCTTTCTCGGAGAA 198
QY 1047 AAAACTGATCTGTGTCAGAGTGTATTTGGGGAGAGGCTTGAAGTGAAGACTGGACAA 1106
DB 199 AAAGCTCTAC----- 209
QY 1107 GCATATCAACTCTGATTTATTATACATCTTTGCTCCAGGAGCAACCAAGATATATGTC 1166
DB 210 -----TCATCAGAGGTATCCAGATATATC 235
QY 1167 TTCTGACAAAGAGAGGCTATACCTTAGAGGTTATCCGAAGCGCTGAGAGAA 1226
DB 236 TTCTGACAAAGAGAGGCTATACCTTAGAGGTTATCCGAAGCGCTGAGAGAA 295
QY 1227 GTCCGAGACCCCTCATGAGATTAATCTGGAATCTGTGAGTGGGCTTTATCTGCTGAG 1286
DB 296 TTTCGAGGCTCATGAGGCTGCTGCTTCAATCTGTGAGTGAAGCGCTTATCTGCTGGA 355
QY 1287 TCTTCTGCGCTCATATCA 1305
DB 356 TCTTCTCAAGCTCTACATCA 374

RESULT 13
ABX41681
ID ABX41681 standard; cDNA; 383 BP.
AC ABX41681;
XX
XX 20-FEB-2003 (first entry)
DT
XX
XX Bovine EST associated with lactation/muscle/fat deposition #6846.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
XX muscle deposition; fat deposition; genome mapping; gene identification;

QY 1288 CTTCTCGGCTCATATCATGAGGAGGCGCTTCTGGG 1327
 DB 476 CTTCCAAGCTCTACGTACATCAGAGGCGGCTTGGTG 515

RESULT 10
 ID ABT09465 standard; DNA; 689 BP.
 XX ABT09465;

XX ABT09465;
 DT 05-DEC-2002 (first entry)

XX Phase-1 Rat CT gene SEQ ID No 553.

XX Rat; toxicity study; rat toxic response gene; toxicological response;
 KM drug development; phase-1 rat CT gene; ds.

XX Ratus sp.

XX MO20026682-A2.

XX 29-AUG-2002.

XX 29-JAN-2002; 2002MO-US02935.

XX 29-JAN-2001; 2001US-264933P.

XX 26-JUL-2001; 2001US-308161P.

XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.

XX Farria G, Hicken SH, Farr SB;

XX WPI; 2002-674961/72.

XX Evaluating the toxicity of an agent, useful in drug development or in
 PT determining toxicological responses to a new drug, by determining the
 PT expression of rat toxicologically relevant genes in the test animal in
 PT response to the test agent -

XX Disclosure; Page 241; 388pp; English.

XX The invention relates to a method used for evaluating the toxicity of an
 CC agent comprising determining the expression of a rat toxic response
 CC gene(s) in the test animal in response to the agent. The method is useful
 CC in drug development, particularly for conducting toxicity studies and
 CC analysis before a new drug or compound is approved for human consumption
 CC or use. The method is also useful in determining toxicological responses
 CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
 CC gene of the invention.

XX Sequence 689 BP; 141 A; 185 C; 190 G; 172 T; 1 other;

Query Match 3.6%; Score 136; DB 24; Length 689;
 Best Local Similarity 65.5%; Pred. No. 5.5e-31;
 Matches 262; Conservative 0; Mismatches 55; Indels 83; Gaps 1;

QY 928 TCACAGGAGGACCACTACTGCGCTGAGACACCGCGGAGTGGCTGCAATAGTGGCC 987

DB 121 TTCAAGTGGCTCCCACTACGCGCTGAGCTCCAGCCGTATGAGTGGCATAGTGGCC 180

QY 988 ATTGCTCATAGTGGCCCGGAGGCTCTTCAAGAGTGAATGCTTTCTGGAGAA 1047

DB 181 ATTGCTCATAGTGGCCCGGAGGCTCTTCAAGAGTGAATGCTTTCTGGAGAG 240

QY 1048 AAATCTATCTGCTGCAAGTGTGTATTGGGGAGAGGCTTGAAGTAGAGCTGGAGCAAG 1107

DB 241 AAAGTCTATCTGA----- 253

QY 1108 CATATCAACTCTATATTATTAACATCCTTTGCTCCAGAGGAGCCAGGATATATCT 1167

DB 254 -----TCAGAGGAGCACTCAAGATATATCT 277

QY 1168 TCCTGACAAAGGAGGCTATACCTAGTAGAGCGTTATCCGAGCGGCTGAGAGAAAG 1227
 DB 278 TCCTGACAAAGGAGGCAATTAACCTAGTAGAGCGTTATCCAGCGGCTGAGAGAAAG 337

QY 1228 TCGGAGCCCTCATGGAATTATCTGAGCTGTGTGATGCGGCTTTATCTGCGGT 1287

DB 338 TTGGAGAGCCCTCCCGGAGATCAGCTTGATACCATATGATCAGGCTTTCTCGCTGGTT 397

QY 1288 CTTCTCGGCTCATATCATGAGGAGGCGCTTCTGGG 1327

DB 398 CTTCCAAGCTCTACGTACATCAGAGGCGGCTTGGTG 437

RESULT 11

ID ABX45158 standard; cDNA; 420 BP.

XX ABX45158;

XX 21-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #10323.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 KM muscle deposition; fat deposition; genome mapping; gene identification;
 KM gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137139-A1.

XX 26-SEP-2002.

XX 24-SEP-2001; 2001US-0960352.

XX 12-JAN-1999; 99US-115707P.

XX 11-JAN-2000; 2000US-0480902.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-110599/10.

PS Claim 2; SEQ ID No 10323; 245bp; English.

XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 1512 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are; (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 1512 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome

PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2; SEQ ID No 7040; 245bp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMPD), derived
CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMPD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMPD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139.
XX
XX Sequence 408 BP; 95 A; 109 C; 111 G; 93 T; 0 other;
SQ
Query Match 4.5%; Score 167; DB 25; Length 408;
Best Local Similarity 79.8%; Pred. No. 9e-41;
Matches 197; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1479 CAGACGCGCGCTGTGTGCTGACCTGAAGTCAGAGGCCCAAGCAGATGACAGAGC 1538
DB 162 CAGGCGAAGAGCTGTGAGAGGCTGACCTTAAGCTTAAGAGCTCAAGCCACGTGACAGAGC 221
QY 1539 TTCCTTGGCCCCATGAGAGAGTGAAGAGCTTGTGTATGAGAAAGTCCCTTGGCCCTTA 1598
DB 222 TTCCTTGGCTCCATGAGAAAGTGAAGAGGCTGTGTATGACAGAGAGTCTCTGGGCCCTC 281
QY 1599 ACTCATGTTCCGCAATGTCCTGGCTTGTATCTCATCTGATGCTCCCAATTGTGATGCT 1658
DB 282 ACTCGGTCTTGTGCAATGTCTGAGGCTGTGATCTGTGCAAGGCCCAATCTGTACTGCT 341
QY 1659 ACAGATATGAGGAAGATGATGACAGCCCAAGGCCCTTCCGCAAGCCCAAGATGACCA 1718
DB 342 ACAGATATGAGGAAGATGATGACAGCCCAAGGCCCTTCCGCAAGGCCCAAGATGACCA 401
QY 1719 GTCTCTCT 1725
DB 402 GCCTCTCT 408
RESULT 9
ABT08987
ID ABT08987 standard; DNA; 631 BP.

XX
XX ABT08987;
AC
XX
XX 05-DEC-2002 (first entry)
DT
XX
XX Phase-1 Rat CT gene SEQ ID No 75.
DE
XX
XX Rat; toxicity study; rat toxic response gene; toxicological response;
KM drug development; phase-1 rat CT gene; ds.
XX
XX Rattus sp.
OS
XX
XX W0200266682-A2.
PN
XX
XX 29-AUG-2002.
PD
XX
XX 29-JAN-2002; 2002MO-US02935.
PF
XX
XX 29-JAN-2001; 2001US-264933P.
PR
XX
XX 26-JUL-2001; 2001US-308161P.
PR
XX
XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
PA
XX
XX Faris G, Hicken SH, Farr SB;
PI
XX
XX WPI; 2002-674961/72.
DR
XX
XX
XX Evaluating the toxicity of an agent, useful in drug development or in
PT determining toxicological responses to a new drug, by determining the
PT expression of rat toxicologically relevant genes in the test animal in
PT response to the test agent
XX
XX
XX Disclosure; Page 124; 388bp; English.
PS
XX
XX The invention relates to a method used for evaluating the toxicity of an
CC agent comprising determining the expression of a rat toxic response
CC gene(s) in the test animal in response to the agent. The method is useful
CC in drug development, particularly for conducting toxicity studies and
CC analysis before a new drug or compound is approved for human consumption
CC or use. The method is also useful in determining toxicological responses
CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC gene of the invention.
XX
XX Sequence 631 BP; 139 A; 168 C; 169 G; 154 T; 1 other;
SQ
Query Match 3.6%; Score 136; DB 24; Length 631;
Best Local Similarity 65.5%; Pred. No. 5.2e-31;
Matches 262; Conservative 0; Mismatches 55; Indels 83; Gaps 1;
QY 928 TCACAGGAGACCCACTACTGAGGCTGTGACACACGAGCGGATGGCTGATGCTGGCCC 987
DB 199 TTGAGTGGCTCCCACTACTGAGGCTGTGACCTCAGCGGATGGCTGATGCTGGCCC 258
QY 988 ATTGCTCATGATGAGCCCGAGGCTCTTCAAGAGTGTGATGCTGCTTTCTGGAGAA 1047
DB 259 ATTGCTCATGATGAGCCCGAGGCTCTTCAAGAGTGTGATGCTGCTTTCTGGAGTGG 318
QY 1048 AAATCTATCTGTGCTCAGGTGTGTATGGGGAGAGGCTTGAAGTGAAGACTGGGACAA 1107
DB 319 AAAGTCTATCTGA----- 331
QY 1108 CATATCAACTCTGATTTATTATACATCTTTGCTCCAGGGGACCCAGGTATATGCT 1167
DB 332 -----TCAGGGCACTCAAGTATATGCT 355
QY 1168 TCTTGACAAAGGAGGCTATACCTAGTAAGCGGTTATCCGAAGCGCTGAGAAAGAA 1227
DB 356 TCTTGACAAAGGAGGCAATTAACCTAGTAAGGTTTATCCAAAGCGCTGAGAAAGAAC 415
QY 1228 TCGGAGCCCTCATGAGGATTTATCTGAGCTGTGTGATCGGCTTTATCTGCCCTGGGT 1287
DB 416 TTGGAGGCCCTCCGGGATCAGCTTGATATACATATAGATGACACTTTTCTGCCCTGGTT 475

QY 1479 CAGGACGGCGGCTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAGCTGACAGAGC 1538
 CC |||||
 DB 366 CAGGCCAAGAGCTGTGAGGCTGACCTGAACTTAGAGGCTCAAGCCAGTGAACAGAGA 307
 CC |||||
 QY 1539 TTCTCTGGCCCAATGAGAGAGTACGAGACCTTGTGTATGAGAAAGTCCCTTGCCCTTA 1598
 CC |||||
 DB 306 TTCTCTGGCTCCATAGAGAAAGTGAAGGGCCCTGTGTATGAGAAAGTATCTGGGCCCC 247
 CC |||||
 QY 1599 ACTCATGTTCCGCAATGCTCCGCGCTTGTACTCATCCATGGTCCCAATTGTGATGCT 1658
 CC |||||
 DB 246 ACTCGATGTGTGCAATGTGTGCTGTGTACTCTGTGCAAGGCCCAATCTGTACTGCT 187
 CC |||||
 QY 1659 ACAGATGATGTGAGAACTGAAATGCAAGCCAGGCGCTTCCGCAACCCAGATGTGACCA 1718
 CC |||||
 DB 186 ACAAGATGTGTGAGAAATGTGACACCAAGGACCTTCCGCAAGGCCCAAGATGACCA 127
 CC |||||
 QY 1719 GTCTCTGGCGCTGACCTGACCTGAGGGGCTTGTGACATGAGTGTGCGCTTGCCCACTC 1778
 CC |||||
 DB 126 GCCTCTGGGCTGCGCTCCCAACAGACCTCTGATGCGAGTTTGGCTCAGATCTCTCTC 67
 CC |||||
 QY 1779 CTAG-TTCTCATATTAAGACAGATTGCTTCTGCTT 1816
 CC |||||
 DB 66 CCAGTTTCATCAATTAAGCCAGAGTGTCTTCACT 28
 CC |||||

RESULT 7
 ABX46151
 ID ABX46151 standard; cDNA; 387 BP.
 AC ABX46151;
 XX
 DT 21-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #11316.
 XX
 KM Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KM muscle deposition; fat deposition; genome mapping; gene identification;
 KM gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-0960352.
 XX
 PR 12-JAN-1999; 99US-115707P.
 PR 11-JAN-2000; 2000US-0480902.
 XX
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 DR WPI; 2003-110599/10.
 XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX
 PS Claim 2; SEQ ID No 11316; 245bp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived
 CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are; (1) a transformed cell having a nucleic acid
 CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-

CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMFD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMFD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.
 XX
 SQ Sequence 387 BP; 86 A; 108 C; 107 G; 86 T; 0 other;

Query Match 4.8%; Score 180.6; DB 25; Length 387;
 Best Local Similarity 78.5%; Pred. No. 5.2e-45;
 Matches 216; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1479 CAGGACGGCGGCTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAGCTGACAGAGC 1538
 CC |||||
 DB 111 CAGGCCAAGAGCTGTGAGGCTGACCTGAAGCTTAGAGGCTCAAGCCAGTGAACAGAGA 170
 CC |||||
 QY 1539 TTCTCTGGCCCAATGAGAGAGTACGAGACCTTGTGTATGAGAAAGTCCCTTGCCCTTA 1598
 CC |||||
 DB 171 TTCTCTGGCTCCATAGAGAAAGTGAAGGGCCCTGTGTATGAGAAAGTATCTGGGCCCC 230
 CC |||||
 QY 1599 ACTCATGTTCCGCAATGCTCCGCGCTTGTACTCATCCATGGTCCCAATTGTGATGCT 1658
 CC |||||
 DB 231 ACTCGATGTGTGCAATGTGTGCTGTGTACTCTGTGCAAGGCCCAATCTGTACTGCT 230
 CC |||||
 QY 1659 ACAGATGATGTGAGAACTGAAATGCAAGCCAGGCGCTTCCGCAACCCAGATGTGACCA 1718
 CC |||||
 DB 291 ACAAGATGTGTGAGAAATGTGACCAAGGACCTTCCGCAAGGCCCAAGATGACCA 350
 CC |||||
 QY 1719 GTCTCTGGCGCTGACCTGACCTGAGGGGCTTCTGTA 1753
 CC |||||
 DB 351 GCCTCTGGGCTGCGCTCCCAACAGACCTCTCTGA 385
 CC |||||

RESULT 8
 ABX41875
 ID ABX41875 standard; cDNA; 408 BP.
 AC ABX41875;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #7040.
 XX
 KM Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KM muscle deposition; fat deposition; genome mapping; gene identification;
 KM gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-0960352.
 XX
 PR 12-JAN-1999; 99US-115707P.
 PR 11-JAN-2000; 2000US-0480902.
 XX

AC AAK53438;
 XX 16-NOV-2001 (first entry)
 XX
 DE Murine transport and binding associated protein encoding cDNA SEQ. ID 3.
 XX
 KM Murine; liver; gene library; amino acid synthesis; binding protein;
 KM cell metabolism; energy metabolism; fatty acid metabolism; synthesis;
 KM phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;
 KM replication; transcription; translation; transport protein; ss.
 XX
 OS Mus musculus.
 XX DE20103510-U1.
 XX 07-JUN-2001.
 XX
 XX 28-FEB-2001; 2001DE-2003510.
 XX
 XX 02-DEC-1999; 99DE-1058160.
 XX
 XX (LION-) LION BIOSCIENCE AG.
 XX
 XX WPI; 2001-368570/39.
 XX
 PT Gene library containing sequences with specific 3'-ends and no polyA
 PT tail, encoding proteins involved in a wide range of cellular processes
 PT
 PS Claim 15; Page 31; 251pp; German.
 XX
 CC This invention describes a novel gene library (A) comprises a gene
 CC sequence (or its part) encoding a protein involved in amino acid
 CC synthesis, cellular/energy metabolism, metabolism of
 CC fatty acids/phospholipids, synthesis or breakdown of
 CC purines/pyrimidines/nucleosides/nucleotides, DNA
 CC replication/transcription/translation, or is a transport/binding protein.
 CC (A) are produced that correspond to the 3'-end of mRNA but without the
 CC polyA tail. They can be prepared more efficiently and with less effort
 CC than conventional libraries. AAK53436-AAK54275 represent fragments of the
 CC gene library described in the method of the invention.
 CC
 SQ Sequence 339 BP; 83 A; 81 C; 90 G; 85 T; 0 other;
 CC
 CC Query Match 5.5%; Score 204; DB 22; Length 339;
 CC Best Local Similarity 77.4%; Pred. No. 2.5e-52;
 CC Matches 263; Conservative 0; Mismatches 70; Indels 7; Gaps 1;
 CC
 QY 1475 TCCCAAGACGCGGCTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAGTGGACA 1534
 DB 339 TCAATCAGACGCGGCTGTGTGCTGACCTGAAGTCAGAGCCCAAGCCAGTGGACA 280
 QY 1535 GAGCTTCTTGGCCCATGAGAAAGTACGAGAGCTTGTATGAAAAGTCCCTTGGC 1594
 DB 279 GAGGTTCCGCGCCCATGAGAAAGTACGAGAGCTTGTATGAAAAGTCCCTTGGC 220
 QY 1595 CCTAATCATGTTCCCGCATGTGTCCGCTTGAATCTATTCATGATGCTCCATTTGTAC 1654
 DB 219 CCAACACAGATGTTCTTCATATGATGTTCAAGCTTATTCATATGAGCCCAATTTGTAC 160
 QY 1655 TGTATACAGATGATGAGAAAGTAAAGTACGAGAGCTTGTATGAAAAGTCCCTTGGC 1714
 DB 159 TGTATACAGATGATGAGAAAGTAAAGTACGAGAGCTTGTATGAAAAGTCCCTTGGC 100
 QY 1715 ACCAGTCTCTGCGCTGACTCACTGAGGGGCTTGTGATGATGATGCTGCGCTGGCCCA 1774
 DB 99 AACAGATCTCTGCGCTGACTCACTGAGGGGCTTGTGATGATGATGCTGCGCTGGCCCA 47
 QY 1775 CCTCTAGTTCCTCATATTAAGACAGATGCTTCTTCCG 1814
 DB 46 TCTCATTTCATCTCATATTAAGACAGATGCTTCTTCCG 7

RESULT 6
 ABX41415/C
 ID ABX41415 standard; cDNA; 400 BP.
 XX
 AC ABX41415;
 XX
 XX 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #6580.
 XX
 KM Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 KM muscle deposition; fat deposition; genome mapping; gene identification;
 KM gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 XX US2002137139-A1.
 XX
 XX 26-SEP-2002.
 XX
 XX 24-SEP-2001; 2001US-0960352.
 XX
 XX 12-JAN-1999; 99US-115707P.
 XX
 XX 11-JAN-2000; 2000US-0480902.
 XX
 PA (BYATT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 DR WPI; 2003-110599/10.
 XX
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX
 XX Claim 2; SEQ ID No 6580; 245bp; English.
 PS
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are; (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridization between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139.
 CC
 SQ Sequence 400 BP; 86 A; 106 C; 115 G; 93 T; 0 other;
 CC
 CC Query Match 5.3%; Score 197.4; DB 25; Length 400;
 CC Best Local Similarity 75.8%; Pred. No. 3.2e-50;
 CC Matches 257; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

Db 292 ACTGATGTTCCGCAATGATGCGGGCTTGACCTATCCATG-CCCAATTTGTACTGCT 234
 QY 1659 ACGTGTATGTGGAAGAACTGATGACAGCCAGCCCTTCCGCAACCCGAGATGTGACCA 1718
 Db 233 AATGATATGTGGAGAACTGAAATGCA-NCAAGGCCCTTCGCAACCCGAGATGTGACCA 175
 QY 1719 GTTCCGGGCTGACCTGACCTGAGGGGCTTTCGACATGATGCTGGGCTGGCCCACTTC 1778
 Db 174 GTTCTCGGCTGACCTGACCTGAGGGGCTTTCGACATGATGCTGGGCTGGCCCACTTC 118
 QY 1779 CTAGTTCCTATATATGAAGACAGATGCTTCTTCTGCTTCTGACGAGGGCTTTCGAC 1838
 Db 117 CTAG-TCTCATATATGAAGACAGATGCTTCTT-GCTTCTGACGAG-GGAGCTTTCGAC 61
 QY 1839 TGAGTCTGGGCTTGGCCCACTTCCCAAGTTTCTATATATGAAGACAGATGCTTTCGAC 1898
 Db 60 TGAGTCTGGGCTTGGCCCACTTCCCAAGTTTCTATATATGAAGACAGATGCTTTCGAC 3
 QY 1899 TT 1900
 Db 2 TT 1
 RESULT 4
 AB283572
 ID AB283572 standard; cDNA; 690 BP.
 AC AB283572;
 XX
 DT 14-MAY-2003 (first entry)
 XX
 DE Toxicologically relevant human nucleotide sequence #731.
 XX
 KW Toxicologically relevant gene; toxicological response; gene; ss.
 OS Homo sapiens.
 XX
 PN MO2003016500-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 16-AUG-2002; 2002MO-US26514.
 XX
 PR 16-AUG-2001; 2001US-313080P.
 PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
 XX
 PI Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeisler K;
 PI Allen P;
 XX
 DR WPI; 2003-268322/26.
 XX
 PT Determining a toxicological response to an agent, useful for screening
 PT of drugs, comprising comparing the expression profile of one or more
 PT human toxic response genes to a reference gene expression profile
 PT indicative of toxicity -
 XX
 PS Claim 1, Page 241, 455pp; English.
 XX
 CC The present invention describes a method (M1) for determining a
 CC toxicological response to an agent, which comprises comparing the
 CC expression profile of one or more human toxic response genes to a
 CC reference gene expression profile indicative of toxicity, and so
 CC determining the presence of a toxic response to the agent. Also
 CC described: (1) an array comprising one or more polynucleotides selected
 CC from the genes corresponding to the partial sequences given in AB282842
 CC to AB284764, or their fragments of at least 20 nucleotides, or
 CC homologues; and (2) determining if a gene putatively identified to be a
 CC toxic response gene plays a role on toxic response pathways by
 CC determining the expression profile of the gene after exposure of cells
 CC or a human subject to a known toxic pharmaceutical or industrial agent,
 CC comprising: (a) exposing cells to an agent or isolating cells from a
 CC human subject who was exposed to an agent; (b) obtaining the test gene

CC expression profile for a putatively identified toxic response gene after
 CC exposure to a known toxic pharmaceutical or industrial agent; and
 CC (c) comparing the test profile to the expression profile of a gene with
 CC a similar function or comparing the test profile to the expression
 CC profile of that gene after exposure to other known toxic compounds. The
 CC methods are useful for predicting and determining toxicological responses
 CC on a cellular, organ or system level. The arrays comprising the human
 CC genes are useful for toxicological screening of drugs, pharmaceutical
 CC compounds and chemicals.
 XX
 SQ Sequence 690 BP; 147 A; 181 C; 196 G; 164 T; 2 other;
 Query Match 6.2%; Score 232.6; DB 25; Length 690;
 Best Local Similarity 66.4%; Pred. No. 5.4e-61;
 Matches 508; Conservative 0; Mismatches 4; Indels 253; Gaps 2;
 QY 927 TTCCAGAGGAGCCCACTACTGCGCTGTGACACCAAGCGGATGCTGCATATGCTGCC 986
 Db 71 TTTCAGTGGAGCCCACTACTGCGCTGTGACACCAAGCGGATGCTGCATATGCTGCC 130
 QY 987 CATGCTCATCAATGCGCCCAAGGCTCTTACAGAGTGAATGCTGCTTTCCTGGGAGA 1046
 Db 131 CATGCTCATCAATGCGCCCAAGGCTCTTACAGAGTGAATGCTGCTTTCCTGGGAGA 190
 QY 1047 AAAACTCTATCTGTCACAGTGTGATTTGGGGAGAGGCTTAGAGACTGGAGCAA 1106
 Db 191 AAAACTCTATCTG----- 204
 QY 1107 GCATATCAACTCTGTATTTATTATACCATCTTTGCTTCAGAGGACCCAGGTATATGTC 1166
 Db 205 -----TCAGGGGACCCAGGTATATGTC 227
 QY 1167 TTCTCGAAGAGAGGCTTATACCTTATAGAGGCTTATCCGAGGCGCTGAGAGAGAA 1226
 Db 228 TTCTCGAAGAGAGGCTTATACCTTATAGAGGCTTATCCGAGGCGCTGAGAGAGAA 287
 QY 1227 GTGCGAGACCCCTCATGAGATTAATCTGACTCTGTGATGCGGCTTTATCTGCGG 1286
 Db 288 GTGCGAGACCCCTCATGAGATTAATCTGACTCTGTGATGCGGCTTTATCTGCGG 347
 QY 1287 TCTTCTCGGCTCCATATATGAGGAGTGAAGGCTTCTGAGTCTTAGAGGCGAGCTTGT 1346
 Db 348 TCTTCTCGGCTCCATATATG----- 369
 QY 1347 TCTGTAAGCTGTCTGTGAGATAGATCCCAAGGCGCATGAGAGGCTTAGAGATC 1406
 Db 370 ----- 369
 QY 1407 CCCAGGCGATGAGAGGCTTAGAGATCCCATGATGAGAGCCATGTATGTTTG 1466
 Db 370 ----- 369
 QY 1467 GTGCTTTCCTCCAGAGAGCGGCTGTGTGCTGTGACCTGAAGTGAAGGCCCAAGCA 1526
 Db 370 -----CAGAGCGGCGCTGTGTGCTGTGACCTGAAGTGAAGGCCCAAGCA 417
 QY 1527 CGTGAAGAGAGCTTCTTGGCCCCCATGAGAGGTATACCGAGGCTTGTATGGAAGT 1586
 Db 418 CGTGAAGAGAGCTTCTTGGCCCCCATGAGAGGTATACCGAGGCTTGTATGGAAGT 477
 QY 1587 CCTTGGCCCTTACTATGTTCCGCAATGATGCTCCGCTTGTATCTATCATCATGCTCCCA 1646
 Db 478 CCTTGGCCCTTACTATGTTCCGCAATGATGCTCCGCTTGTATCTATCATCATGCTCCCA 537
 QY 1647 ATTTGTAAGTCTACAGTATGAGAGAACTGAATGACGCAAG 1691
 Db 538 ATTTGTAAGTCTACAGTATGAGAGAACTGAATGACGCAAG 582
 RESULT 5
 AAK53438/c
 ID AAK53438 standard; cDNA; 339 BP.
 XX

CC of the invention.

XX Sequence 1610 BP; 355 A; 439 C; 444 G; 372 T; 0 other;

Query Match 11.9%; Score 446; DB 22; Length 1610;
Best Local Similarity 73.7%; Pred. No. 4.8e-127;
Matches 722; Conservative 0; Mismatches 5; Indels 253; Gaps 2;

```

QY 928 TCACAGGAGGACCCACTAGTGGGCTCTGACACACAGCGGGATGCGGTGATGAGTGGCCC 987
DB 880 TTCAGTGGACCCACTAGTGGGCTCTGACACACAGCGGGATGCGGTGATGAGTGGCCC 939
QY 988 ATTGCTCATAGTGGCCGACAGGCTCTTACAGCATGATGCTGCTTTCTTCTGGAGAA 1047
DB 940 ATTGCTCATAGTGGCCGACAGGCTCTTACAGCATGATGCTGCTTTCTTCTGGAGAA 999
QY 1048 AAATCTATCTGTGTCAGGTTGTATTTGGGGAGAGGCTTGAAGTAAAGCTGGGCAAG 1107
DB 1000 AAATCTATCTGTGTCAGGTTGTATTTGGGGAGAGGCTTGAAGTAAAGCTGGGCAAG 1012
QY 1108 CATATCAACTCTGTATTTATTATACATCCTTTGCTCCAGGGACCCAGTATATGCT 1167
DB 1013 -----TCCAGGGACCCAGTATATGCT 1036
QY 1168 TCCTGACAAAGGAGGCTATACCTTAGTAAGCGGTTATCCGAAGCGGCTGAGAAAGAA 1227
DB 1037 TCCTGACAAAGGAGGCTATACCTTAGTAAGCGGTTATCCGAAGCGGCTGAGAAAGAA 1096
QY 1228 TCGGAGCCCTCATAGGATTTATCTGAGCTCTGTGATGCGGCTTTATCTGCGTGGGT 1287
DB 1097 TCGGAGCCCTCATAGGATTTATCTGAGCTCTGTGATGCGGCTTTATCTGCGTGGGT 1156
QY 1288 CTTTCGSGCTCATATCATGATGAGGAGGAGGCTTGTGGTGTCTTGAAGGACGCTTGT 1347
DB 1157 CTTTCGSGCTCATATCATGATGAGGAGGAGGCTTGTGGTGTCTTGAAGGACGCTTGT 1177
QY 1348 CTGCTACTGTCTGTGAGCATAGATCCCAACAGAGGATGAGAGGCTAGTCAAGATCC 1407
DB 1178 ----- 1177
QY 1408 CCAAGGCGATGAGAGGCTAGGTCAGAGATCCCATGATCCCATGAGAGGCTAGTGTGG 1467
DB 1178 ----- 1177
QY 1468 TGGCTTCTCCAGAGCGGCGGTGTGAGTGGCTGACCTGAAGTCAAGAGCCCAAGCCAC 1527
DB 1178 -----CAGAGCGGCGGTGTGAGTGGCTGACCTGAAGTCAAGAGCCCAAGCCAC 1226
QY 1528 GTGACAGAGCTTCTTGGCCCATGAGAGGATGAGAGGCTTGTATGAGAAAGTTC 1587
DB 1227 GTGACAGAGCTTCTTGGCCCATGAGAGGATGAGAGGCTTGTATGAGAAAGTTC 1286
QY 1588 CTTTGGCCCTTAATCTATGTTCCGCAATGATGCTCCGCTTGTACCTGATCCATGCTCCAA 1647
DB 1287 CTTTGGCCCTTAATCTATGTTCCGCAATGATGCTCCGCTTGTACCTGATCCATGCTCCAA 1346
QY 1648 TTGTGACTGCTCAAGTATGAGAGAAATGATGAGAGGAGCCCTTCCGCAAGCCCA 1707
DB 1347 TTGTGACTGCTCAAGTATGAGAGAAATGATGAGAGGAGCCCTTCCGCAAGCCCA 1406
QY 1708 GAATGACAGAGTCTCTGGGCTGACACTGAGAGGGGCTTCTGACATGAGTCTGGCCT 1767
DB 1407 GAATGACAGAGTCTCTGGGCTGACACTGAGAGGGGCTTCTGACATGAGTCTGGCCT 1466
QY 1768 GGGCCCACTCTAGTCTCTCAATATMAAGACAGATTTCTTCCGCTTCTCACTGAGGG 1827
DB 1467 GGGCCCACTCTAGTCTCTCAATATMAAGACAGATTTCTTCCGCTTCTCACTGAGGG 1526
QY 1828 GCGCTTCTGACATGAGTCTGGGCTGGGCGCACTCCCAAGTTTCTCAATATMAAGACAGAT 1887
DB 1527 GCGCTTCTGACATGAGTCTGGGCTGGGCGCACTCCCAAGTTTCTCAATATMAAGACAGAT 1586
QY 1888 TGGTCTTCACTTGAATCAA 1907

```

DB 1587 TGGTCTTCACTTGAATCAA 1606

RESULT 3
ABN96669/c
ID ABN96669 standard; DNA; 504 BP.

XX ABN96669;

DT 13-AUG-2002 (first entry)

DE Gene #3167 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX Homo sapiens.

XX WO200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30589.

XX 02-OCT-2000; 2000US-237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer,
XX hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX liver tissue sample -

XX Claim 1; SEQ ID NO 3167; 298bp; English.

XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumor in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytosolic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 504 BP; 123 A; 119 C; 138 G; 110 T; 14 other;

Query Match 7.1%; Score 265.2; DB 24; Length 504;
Best Local Similarity 93.1%; Pred. No. 3.2e-71;
Matches 393; Conservative 0; Mismatches 16; Indels 13; Gaps 11;

```

QY 1480 AGAGCGGCGGTGTGTGAGCTGAGAGTCAAGAGCCCAAGCCAGTGAAGAGCT 1539
DB 410 AGAGCGGCGGTGTGTGAGCTGAGAGTCAAGAGCCCAAGCCAGTGAAGAGCT 352
QY 1540 TCCCTTGG-CCCAATGAGAAAGTGAAGCGAGCCTTGTATGAGAAAGTCCCTTGGCCCTA 1598
DB 351 TCCCTTGGNCCCAATGAGAAAGTGAAGCGAGCCTTGTATGAGAAAG-NCCCTTGGCCCTA 293
QY 1599 ACTGATGTCGCGCAATGATGCTCCGCGCTTGTACTCATCATGATGCTCCCAATTGTACTGCT 1658

```

PT New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is
 PT associated with a cancer, immunopathology or neuropathology -

XX Claim 1; Page 244; 327bp; English.

CC AAH5161 to AAH5756 represent cell and tissue specific polynucleotide
 CC sequences (I). (I) can have cytostatic, immunomodulatory and
 CC neuroprotective activities, and can be used in gene therapy. (I) and
 CC proteins (II) encoded by them are used in high throughput screening
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
 CC their fragments, immunoglobulins, inhibitors, drug compounds and
 CC pharmaceutical agents. Expression of (I) in a sample indicates the
 CC differentiation of embryonic stem cells into a tissue selected from
 CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
 CC tissues. (I) and (II) are used to produce an expression profile that
 CC defines a metabolic or developmental process, treatment, condition,
 CC disease or disorder. The gene profile can be used for diagnosis,
 CC prognosis or monitoring of treatments and for investigating a
 CC predisposition to a disorder where the gene is associated with a
 CC cancer, immunopathology or neuropathology.

XX Sequence 1571 BP; 344 A; 431 C; 434 G; 362 T; 0 other;

Query Match 11.9%; Score 446; DB 22; Length 1571;
 Best Local Similarity 73.7%; Pred. No. 4.7e-127;
 Matches 722; Conservative 0; Mismatches 5; Indels 253; Gaps 2;

QY 928 TCCAGAGGAGCCACTACTGCGCTGAGACACAGCCGGAGTGGAGCTAGTGGCC 987
 DB 840 TTGAGTGGAGCCACTACTGCGCTGAGACACAGCCGGAGTGGAGCTAGTGGCC 899
 QY 988 ATTGCTCATGATGGCCCGCCAGGGTCTTCAGACAGTGGATGCTCTTTCTGGAGAA 1047
 DB 900 ATTGCTCATGATGGCCCGCCAGGGTCTTCAGACAGTGGATGCTCTTTCTGGAGAA 959
 QY 1048 AACTCTATGTGTCCAGGCTGTATTTGGGGAGAGGCTTAGTAGACTGGACAG 1107
 DB 960 AACTCTATGTGTG----- 972
 QY 1108 CATATCAACTGTGTATTTATATACATCTTGTCTCCAGAGGACCCAGATATATGTCT 1167
 DB 973 -----TCCAGGGACCCAGATATATGTCT 996
 QY 1168 TCCTGACAAAGGAGGCTATACCTTAGTAAGCGTTATCCGAAACGGCTGGAGAGAG 1227
 DB 997 TCCTGACAAAGGAGGCTATACCTTAGTAAGCGTTATCCGAAACGGCTGGAGAGAG 1056
 QY 1228 TCGGAGACCCCTCATGAGATTATCTCTGACCTCTGTGATGGCGCTTTATCTGCCCTGGGT 1287
 DB 1057 TCGGAGACCCCTCATGAGATTATCTCTGACCTCTGTGATGGCGCTTTATCTGCCCTGGGT 1116
 QY 1288 CTCTCGGGCTCATATCATATCATGAGAGTGGGGCTTCTGGGTGTAGAGGGCAGCTTGT 1347
 DB 1117 CTCTCGGGCTCATATCATG----- 1137
 QY 1348 CTGCTACTGTCTGTGGCATATGATCCCAACAGGGCATGAGAGCGCTAGTCAAGATCC 1407
 DB 1138 ----- 1137
 QY 1408 CCAAGGCGATGAGAGGCTTAGTCAAGATCCCAATGACATGAGAGCATGCTATGTTGG 1467
 DB 1138 ----- 1137
 QY 1468 TGCCTTCTCCAGAGCGCGGCTGTGTGCTGAGACCTTAGAGTCAAGAGCCCAAGCCAC 1527
 DB 1138 -----CAGAGCGCGGCTGTGTGCTGAGACCTTAGAGTCAAGAGCCCAAGCCAC 1186
 QY 1528 GTGAGACAGACTTCTTGGCCCAATGAGAGGATGAGAGGAGCTTGTGTATGAGAAAGTC 1587
 DB 1187 GTGAGACAGACTTCTTGGCCCAATGAGAGGATGAGAGGAGCTTGTGTATGAGAAAGTC 1246

QY 1588 CCTGGCCCTAATCATATGTTCCGCAATGATGCCGCTTGTATCCTCATGATGTCGCA 1647
 DB 1247 CCTGGCCCTAATCATATGTTCCGCAATGATGCCGCTTGTATCCTCATGATGTCGCA 1306
 QY 1648 TTTGATGCTCTACAGTATGATGAGAACTGATGAGCCCAAGCCCTTCCGCAACCCA 1707
 DB 1307 TTTGATGCTCTACAGTATGATGAGAACTGATGAGCCCAAGCCCTTCCGCAACCCA 1366
 QY 1708 GAATGACACAGTCTCTGGGCTGCACTCTGAGAGGCTTCTGACATGAGTCTGGCT 1767
 DB 1367 GAATGACACAGTCTCTGGGCTGCACTCTGAGAGGCTTCTGACATGAGTCTGGCT 1426
 QY 1768 GGCCCACTCTAGTCTCTATTAATTAAGACAGATGCTCTGCTCTGCACTGAGGG 1827
 DB 1427 GGCCCACTCTAGTCTCTATTAATTAAGACAGATGCTCTGCTCTGCACTGAGGG 1486
 QY 1828 GCTTCTGACATGATGCTGGCTGGCCCACTCCCACTTTCTCATTAATTAAGACAGAT 1887
 DB 1487 GCTTCTGACATGATGCTGGCTGGCCCACTCCCACTTTCTCATTAATTAAGACAGAT 1546
 QY 1888 TGCTTTCTTCACTTGAATCAA 1907
 DB 1547 TGCTTTCTTCACTTGAATCAA 1566

RESULT 2
 AAH98592
 ID AAH98592 standard; cDNA; 1610 BP.

XX AAH98592;

DT 12-OCT-2001 (first entry)

XX Human EST-derived coding sequence SEQ ID NO: 449.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;

XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;

XX diagnostic; forensic test; gene mapping; genetic disorder;

XX biodiversity; gene therapy; nutrition; ss.

XX Homo sapiens.

XX WO200154477-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02687.

XX 25-JAN-2000; 2000US-0491404.

XX 17-JUL-2000; 2000US-0617746.

XX 03-AUG-2000; 2000US-0631451.

XX 15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSBO INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V,

XX Cao Y, Drmanac RA, Zhang J, Weirman T;

XX WPI; 2001-476164/51.

XX P-PDB; AAM23933.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising

XX antibodies and research use -

XX Claim 1; Page 498; 1275bp; English.

XX The present invention provides the protein and coding sequences of novel

XX proteins from a variety of organisms, including human, dog, cat, horse,

XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea

XX urchin and tomato. These were derived from expressed sequence tags (ESTs)

XX from the organism of interest. They can be used in diagnostics,

XX forensic, gene mapping, identification of mutations, to assess

XX biodiversity and for nutritional purposes. The present sequence is a cDNA

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 06:38:47 ; Search time 586.508 Seconds
(without alignments)
17159.770 Million cell updates/sec

Title: US-09-900-448-3_COPY_10001_13737

Perfect score: 3737
Sequence: 1 cccaagagagagtcagtcac.....accttttctgtgtcttca 3737

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_15Jun03:*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	446	11.9	1571	22	AAH57485	Human liver cell s
2	446	11.9	1571	22	AAH57485	Human liver cell s
3	265.2	7.1	504	24	ABN96659	Gene #167 used to
4	232.6	6.2	690	25	ABK83572	Toxicologically re
5	204	5.5	339	22	AAK53438	Murine transport a
6	197.4	5.3	400	25	ABX41415	Bovine EST associa
7	180.6	4.8	387	25	ABX46151	Bovine EST associa
8	167	4.5	408	25	ABX41875	Bovine EST associa

9	136	3.6	631	24	ABT08987
10	136	3.6	629	25	ABT09445
11	123.6	3.3	480	24	ABX45158
12	118.2	3.2	374	25	ABX40203
13	117.4	3.1	383	25	ABX41681
14	111.2	3.0	288	25	ABX43075
15	106	2.8	397	25	ABX47771
16	105.4	2.8	700	22	AAH92374
17	105.4	2.8	397658	25	ABO83210
18	103.2	2.8	293	25	ABX44025
19	102.2	2.7	343	25	ABX47807
20	101.8	2.7	123219	23	AAH88703
21	101.2	2.7	2674	22	AAH14203
22	101.2	2.7	183999	22	AAH92831
23	100.2	2.7	593	22	AAK07755
24	100.2	2.7	593	22	AAK36622
25	100.2	2.7	593	22	AAK36622
26	100.2	2.7	593	23	ABX33439
27	100.2	2.7	593	23	ABX33439
28	100	2.7	395	21	AAK10672
29	99.2	2.7	9160	24	ABK69800
30	97.6	2.6	9266	24	ABK69899
31	97.6	2.6	9268	24	ABK69898
32	97.6	2.6	12036	22	AAK69341
33	97.4	2.6	1578	24	ABK02154
34	96	2.6	12049	22	AAK69340
35	95.8	2.6	159095	25	ABZ80818
36	95.2	2.5	74822	24	ABT10752
37	95.2	2.5	397658	25	ABO83210
38	94.2	2.5	568	22	ABK63861
39	94.2	2.5	568	22	ABK1041
40	94.2	2.5	568	22	AAK12371
41	94.2	2.5	568	22	AAK38088
42	94.2	2.5	568	22	AAI18862
43	94.2	2.5	568	22	AAI43584
44	94.2	2.5	568	23	ABX37705
45	94.2	2.5	568	24	ABX12100

ALIGNMENTS

RESULT 1
AAH57485 standard; cDNA; 1571 BP.
AAH57485;
10-SEP-2001 (first entry)
Human liver cell specific cDNA sequence SEQ ID NO:325.
Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
lung; liver; uterus; ovary; stomach; kidney; pancreas; se;
metabolic disease; developmental disease; cytostatic; immunomodulatory;
neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
Homo sapiens.
WO200132927-A2.
10-MAY-2001.
02-NOV-2000; 2000MO-US30396.
04-NOV-1999; 99US-0163508.
(INCY-) INCYTB GENOMICS INC.
Sornasse T, Seilhamer JJ, Watson GA,
WPI, 2001-291057/30.

THIS PAGE BLANK (USPTO)

REFERENCE	Direct Submission Submitted (11-SEP-1989) Tatum F.M., National Animal Disease Centre, P O Box 70, Ames IA 50010, U S A 2 (bases 1 to 1486)
AUTHORS	Morgan, W. T., Muster, P., Tatum, F., Kao, S. M., Alam, J. and Smith, A.
TITLE	Identification of the histidine residues of hemopexin that coordinate with heme iron and of a receptor-binding region
JOURNAL	J. Biol. Chem. 268 (9), 6256-6262 (1993)
MEDLINE	93203213
FEATURES	7681064
	Location/Qualifiers

sig_peptide	3. .80			
mat_peptide	78. .1382			
polya_signal	/product="hemopexin"			
	1437. .1442			
BASE COUNT	354 a	381 c	393 g	358 t
ORIGIN				

	Query Match	Best Local Similarity	6.8%; Score 255.6; DB 4; Length 1486;
	Matches	294; Conservative	0; Mismatches 44; Indels 2; Gaps 1;
QY	1479	CAGGACGGCGGCTGTGTGTGAGCTGGAACCTGAAGTCAGAGGCCAAGCCAGTGGACAGAGC	1538
Db	1123	CAGGACGGCGCTGTGTGTGAGCTGGACCTGAAGTCTGAGGCTCAAGCCAGTGGACAGAAC	1182
QY	1539	TTCTTTGGCCCCCATAGAGAAAGTAAACGGAGCCTTGTGTATGGAAAAGTCCCTTGGCCCTA	1598
Db	1183	TTCTTTGGCCCTCAATGAGAAAGGTGAACGGGGGCCCTGTGTATGAAAAGCCCTTGGCCCCA	1242
QY	1599	ACTCATGTTCGGGCAATAGTATCCGGGCTTGTAACCTCATGCATGATGATGCCAATTGTATCGCT	1658
Db	1243	ACTCATGTTCCTACCAAGTGTGTCCCACTTGTAACCTCATGCATGATGATGCCAATTGTATCT	1302
QY	1659	ACAAGTATGTGTGAGAAACTGAATGACAGCCAAAGGCCCTTCGCCAAACCCGAGATGTGACA	1718
Db	1303	ACAAGCAATGTGTGACAAACTCAATCAGCACAAGAAACCTCCCAAGCCCAAGAGATGTGTCA	1362
QY	1719	GTCTTCCTGGGCTGCACTCATGAGGGGCTTTCTGATCATGAGTCTTGAGCCCACTC	1778
Db	1363	GACTCCCTGGGCTGTACTCACTGAAGTGTCTCCGACATGAAATCTGGGCCCACTCCTCTC	1422
QY	1779	C--TAGTTCCTCATATATTAAGAAGATGTGCTTTCTTCGTT	1816
Db	1423	CTGATATTTCTCACAATATTAAGCCAGATTTCTTCTTCACTT	1462


```

RESULT 12
HSHEPEXR      1485 bp      mRNA      linear      PRI 06-JUL-1989
LOCUS          Human mRNA for serum beta-glycoprotein hemopexin.
DEFINITION     X02537.1 GI:32061
ACCESSION      X02537.1 GI:32061
VERSION        glycoprotein; serum glycoprotein.
KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
                Altuda, F., Poli, V., Restagno, G., Argos, P., Cortese, R. and
                Silengo, L.
TITLE          The primary structure of human hemopexin deduced from cDNA
                sequence: evidence for internal, repeating homology
JOURNAL        Nucleic Acids Res. 13 (11), 3841-3859 (1985)
MEDLINE        85242073
PUBMED        2989777
COMMENT        Data kindly reviewed (07-JAN-1986) by L. Silengo.
FEATURES
    source
        1..1485
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        7..1323
        /note="hemopexin; Protein sequence is in conflict with the
        conceptual translation"
        /codon_start=1
        /protein_id="CAA26382.1"
        /db_xref="GI:1335098"
        /translation="PTLPPTSAGNVAABSETRDPVYERCSXGNSFDTATLDDNCTM
        LFFKGFVWKSHEKMBRELISERWKNPSPVDAAFRGSHNSVFLIKDKVWVYPERKCE
        KGPVKLQDEPFIPIPSPLDAAVECHRGECQAEGLVFPQDRFWMDLATGKWKSEWP
        AVNGCSSALRWLGRYYCFQGNQFLRPDPVAGVPPRPDVRDVFMPCEGRGHGHPG
        TGGNTHGHPERWRCSPHVLVSLATSDNMGATYASGTHVMYMDLDSRDGMSWPIAH
        QMPQGSADVDAAPSWMEKLYLVGGTQVYVPLTGTGCTVLYSGVYKREKVPDPIGLI
        DSVDAFRTCPGSRHLHIMAGRRLIMDLDSGAQATTELPWPEKTDGLCNKESLGP
        NCSANGPELVLYLHGNPLCYSDVEKLNAAKALPOPNTSLGCTH"
        1375..1380
        /note="pot. polyA signal"
        1459..1464
        /note="pot. polyA signal"
        1485
        /note="put. polyA site"
    misc_feature
        1485
        /note="put. polyA site"
    polyA_site
        1485
        /note="put. polyA site"
BASE COUNT     323 a      412 c      406 g      344 t
ORIGIN
Query Match    11.5%; Score 431.2; DB 9; Length 1485;
Best Local Similarity 73.6%; Pred. No. 3.8e-124;
Matches 717; Conservative 0; Mismatches 3; Indels 254; Gaps 3;

```

```

QY      1228 TCAGAGACCCCTCATGGGATTTATCTGCACTCTGTGATGCGGCTTTATCTGGCTGGGT 1287
DB      983 TCAGAGACCCCTCATGGGATTTATCTGCACTCTGTGATGCGGCTTTATCTGGCTGGGT 1042
QY      1288 CTTCTGGGCTTCATATCATATGAGGAGGAGGCTTCTGGGCTTTAGAGGAGCTTGGTT 1347
DB      1043 CTTCTGGGCTTCATATCATATGAGGAGGAGGCTTCTGGGCTTTAGAGGAGCTTGGTT 1063
QY      1348 CTGCTACTGTCTGTGGCATTAATCCCAACAGGAGCATGAGAGGCTTAGTCAGATCC 1407
DB      1064 ----- 1063
QY      1408 CCAGGGCATGAGAGGCTTAGTCAGATCCCATGATGAGAGGAGGCTTAGTCAGATCC 1467
DB      1064 ----- 1063
QY      1468 TGCCTTCTCCCAAGAGCGGCGCTGTGCTGAGCTGAAAGTCAGAGCCCAAGCCAC 1527
DB      1064 -----CAGAGCGGCGGCTGTGCTGAGCTGAAAGTCAGAGCCCAAGCCAC 1112
QY      1528 GTGACAGAGCTTCTTGGCCCCCATGAGAGGTAACGAGGCTTGTATGAGAAAGTC 1587
DB      1113 GTGACAGAGCTTCTTGGCCCCCATGAGAGGTAACGAGGCTTGTATGAGAAAGTC 1172
QY      1588 CCTGGCCCTTAACTCATGTTTCCGCAATGATGTCGGGCTTGTACCTCATGATGATCCCA 1647
DB      1173 CCTGGCCCTTAACTCATGTTTCCGCAATGATGTCGGGCTTGTACCTCATGATGATCCCA 1232
QY      1648 TTTGTACTGCTCAAGTATGAGAGAACTGATGAGCCCAAGGCTTCCGCAAGCCCA 1707
DB      1233 TTTGTACTGCTCAAGTATGAGAGAACTGATGAGCCCAAGGCTTCCGCAAGCCCA 1292
QY      1708 GAATGACCAAGTCTTCTGGGCTGCACTCACTGAGGAGGCTTCTGACATGAGTTCGCT 1767
DB      1293 GAATGACCAAGTCTTCTGGGCTGCACTCACTGAGGAGGCTTCTGACATGAGTTCGCT 1351
QY      1768 GGGCCCAAGTCTTCTGGGCTGCACTCACTGAGGAGGCTTCTGACATGAGTTCGCT 1827
DB      1352 GGGCCCAAGTCTTCTGGGCTGCACTCACTGAGGAGGCTTCTGACATGAGTTCGCT 1411
QY      1828 GCGTCTGACATGATGCTGGGCTGAGCCCAAGTCTTCTGATTAATAAGACAGAT 1887
DB      1412 GCGTCTGACATGATGCTGGGCTGAGCCCAAGTCTTCTGATTAATAAGACAGAT 1471
QY      1888 TCGTCTTCTGCTTG 1901
DB      1472 TCGTCTTCTGCTTG 1485

```

```

RESULT 13
G11396/c
LOCUS          G11396      400 bp      DNA      linear      STS 30-MAR-2000
DEFINITION     SHGC-10293 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION      G11396
VERSION        G11396.1 GI:1017488
KEYWORDS       STS.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE          1 (bases 1 to 400)
JOURNAL        Unpublished, Olivier, M., Cox, D.R. (2000)
COMMENT        Unpublished (2000)
                Contact: Michael Olivier, David R. Cox
                Stanford Human Genome Center
                Stanford University School of Medicine
                4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
                Tel: (650) 320-5800
                Fax: (650) 320-5801
                Email: olivier@shgc.stanford.edu

```

Db 1458 GGGCCCACTCCAGTTCCTCATATAAAGACAGATGCTTTCGCTTCTCACTAGGG 1517
QY 1828 GCCTTCGACATAGTCTGGCCCTGGCCCACTCCCACTTTCATATAAAGACAGAT 1887
Db 1518 GCCTTCGACATAGTCTGGCCCTGGCCCACTCCCACTTTCATATAAAGACAGAT 1577
QY 1888 TGCTTCTTCACTTGAATCAA 1907
Db 1578 TGCTTCTTCACTTCAAAAAA 1597
RESULT 11
HUMHXA 1546 bp mRNA linear PRI 08-NOV-1994
LOCUS Huma hemopexin mRNA, 3' end.
DEFINITION J03048.1 GI:184487
ACCESSION J03048.1 GI:184487
KEYWORDS hemopexin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Law, M.L., Cai, G.Y., Hartz, J.A., Jones, C. and Kao, F.T.
TITLE 1 (bases 1 to 1546)
The hemopexin gene maps to the same location as the beta-globin
gene cluster on human chromosome 11
JOURNAL Genomics 3 (1), 48-52 (1988)
MEDLINE 89122012
PUBMED 3220477
COMMENT Original source text: Human, cDNA to mRNA, clone lambda-Hx5.
Draft entry and printed copy of sequence for (Genomics (1988) In
press) kindly provided by
G.-Y. Cai and M.L. Law, 01-June-1988.
FEATURES
source
1.1546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11p15.5-p15.4"
1.1546
/gene="HPX"
/gene="HPX"
/gene="HPX"
/product="HXM mRNA"
<1.1386
/gene="HPX"
/note="hemopexin precursor"
/codon_start=1
/protein_id="AA52704.1"
/db_xref="GI:386789"
/db_xref="GDB:G00-120-054"
/translations="ARVLGAPVALGLMSLWSLAITPLPPTISAHGNVABGETKDPD
VTERCDGWSFDAITLDDNTMLFKGGEFVWKSHKWDRELIISRMKNFSPVDAARQ
GHNSVPLIKGDKYVVPPEKKEGYPLIQDEPGIIPSDAAVECHRESCQEGVLE
FOGDRWFMDLATGTWKERSMPPAVGNCSSALRWLGRYYCQGNQPLRPDVEVPFR
YPRDVEDYFMPGCGRGHGRNGHNSHLSHGBEVMRCSPHVLASLTSDNHGTAF
SCDHVYRLDTSRNGHSMWPIAHOWPGSPSVDAAFSPBEKLYIVQSTQYVETKGGY
TLVSGYFKRLKEKVGTPHGIILDSVDAAFICPSSRLHTIAGRLWMLDKSGAQATW
TELPWHEKVDGALCMKESLSPNSCSANGGLYLIGHPNLYCVSDVEKLNAAKALPOP
QNTVSLGCTH"
<1.66
/gene="HPX"
/note="hemopexin signal peptide"
67.1383
/gene="HPX"
/product="hemopexin"
1380.1462
/note="tandem repeat copy A"
1463.1546
/note="tandem repeat copy B"
BASE COUNT 330 a 427 c 430 g 359 t
ORIGIN Chromosome 11pter-p11.

Query Match 11.9%; Score 443.2; DB 9; Length 1546;
Best Local Similarity 73.7%; Pred. No. 66-128;
Matches 718; Conservative 0; Mismatches 3; Indels 253; Gaps 2;
QY 928 TCACAGGAGACCCACTACTAGGCGCTGACACACAGCCGAGATGGCTGACATAGCTGGCCC 987
Db 826 TTCAGTGGAGACCCACTACTAGGCGCTGACACACAGCCGAGATGGCTGACATAGCTGGCCC 885
QY 988 ATTGCTCATAGTGGCCCGAGAGGTCTTTCAGACATGATGATGCTTTCCTGGGAAGAA 1047
Db 886 ATTGCTCATAGTGGCCCGAGAGGTCTTTCAGACATGATGATGCTTTCCTGGGAAGAA 945
QY 1048 AAATCTATCTGGTCCAGGTGTGTATTGGGGAGAGAGCTTGAAGTGAAGACTGGACAAG 1107
Db 946 AAATCTATCTGG----- 958
QY 1108 CATATCAACTGTGATTTATTAACATCCTTGTCTCCAGAGGACCCAGATATATGCT 1167
Db 959 -----TCCAGGGCACCCAGGTATATGCT 982
QY 1168 TCCTGACAAAGGAGGCTATACCTTAGTAAGCGGTATTCGAAAGCGGCTGGAGAGAG 1227
Db 983 TCCTGACAAAGGAGGCTATACCTTAGTAAGCGGTATTCGAAAGCGGCTGGAGAGAG 1042
QY 1228 TGGGAGCCCTCATGAGGATTAATCTGAGCTCTGTGATGAGCGGCTTATCTGCGCTGGGT 1287
Db 1043 TGGGAGCCCTCATGAGGATTAATCTGAGCTCTGTGATGAGCGGCTTATCTGCGCTGGGT 1102
QY 1288 CTTCCTGGCTCCATATCATATGAGGAGGTGAGGGGCTTCTGGGTGCTTAAGGGCAGCTTGT 1347
Db 1103 CTTCCTGGCTCCATATCATATG----- 1123
QY 1348 CTGCTACTGTGTGTGAGCATATATCCCAACAGGGGATGAGAGGCTTAGATGAGATCC 1407
Db 1124 ----- 1123
QY 1408 CCAGGCGATGAGAGGCTTAGATGAGATCCCATGATGACATGAGAGGCTATGTTGG 1467
Db 1124 ----- 1123
QY 1468 TGCTTCTCCCAAGAGCGGCGGTGTGTGCTGACCTGAAAGTCAAGAGCCCAAGCCAC 1527
Db 1124 -----CAGAGCGGCGGTGTGTGCTGACCTGAAAGTCAAGAGCCCAAGCCAC 1172
QY 1528 GTGAGACAGGCTTCTTGGCCCCCATGAGAGGTGAGCGAGCTTGTGTATGAGAAAGTC 1587
Db 1173 GTGAGACAGGCTTCTTGGCCCCCATGAGAGGTGAGCGAGCTTGTGTATGAGAAAGTC 1232
QY 1588 CCTTGGCCCTTAATCTATGTTCCGCAATGATGATCCGAGCTTGTATCTCATCATGATGATCCCA 1647
Db 1233 CCTTGGCCCTTAATCTATGTTCCGCAATGATGATCCGAGCTTGTATCTCATCATGATGATCCCA 1292
QY 1648 TTTGTACTGCTACAGTATGTGAGAGAACTGAATGAGCAAGGCCCTTCCGCAAGCCCA 1707
Db 1293 TTTGTACTGCTACAGTATGTGAGAGAACTGAATGAGCAAGGCCCTTCCGCAAGCCCA 1352
QY 1708 GAATGTGACAGTCTCTGGGCTGACCTCACTGAGGGGCTTGTGACATGATGATGATGCT 1767
Db 1353 GAATGTGACAGTCTCTGGGCTGACCTCACTGAGGGGCTTGTGACATGATGATGATGCTG 1412
QY 1768 GGGCCCACTCCAGTTCCTCATATAAAGACAGATGCTTTCGCTTCTCACTAGAGG 1827
Db 1413 GGGCCCACTCCAGTTCCTCATATAAAGACAGATGCTTTCGCTTCTCACTAGAGG 1472
QY 1828 GCCTTCTGACATGATGCTGGCCCTGGCCCACTCCCACTTTCATATAAAGACAGAT 1887
Db 1473 GCCTTCTGACATGATGCTGGCCCTGGCCCACTCCCACTTTCATATAAAGACAGAT 1532
QY 1888 TGCTTCTTCACTTG 1901
Db 1533 TGCTTCTTCACTTG 1546

Db 207456 AT 207455

RESULT 10
BC005395
LOCUS
DEFINITION Homo sapiens cDNA clone MGC:12533 IMAGE:4071616, complete cds.
ACCESSION BC005395
VERSION BC005395.1 GI:13529280
KEYWORDS MGC.
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalski, U., Smailus, D.E., Schnerk, A., Schein, J.E., Jones, S.U. and Marra, M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
12477932
2 (bases 1 to 1621)
Strausberg, R.
Direct Submission
Submitted (27-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdcpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAL Plate: 16 Row: h Column: 22.
Location/Qualifiers
1..1621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12533 IMAGE:4071616"
/tissue_type="Liver"
/clone_id="NIH_MGC_76"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
62..826

```

/codon_start=1
/product="Unknown (protein for MGC:12533)"
/protein_id="AA05395.1"
/db_xref="GI:13529281"
/translations="MARVLGAPVALGLMSLSLATATLPPTPSAGNNVABGETKDP
DYTERCSDEWSFDATLDDNGTMLPFKGEFVWSHMDRELISERMKNFSPVDAAR
OGHNSVFLIKGDVWVYPPBKKEKGPXKLQDEFPPIPSFLDAVECHRECAEGVL
FPGDDEWFMDLATGTMKERSMPAVGNCSALMLGRYYCFQGNQFLRFPVGEVVP
RYPRIYVDYFMPGGRGHGRNGTGNISICAVAH."
BASE COUNT      370 a      438 c      441 g      372 t
ORIGIN
Query Match      11.9%; Score 444.4; DB 9; Length 1621;
Best Local Similarity 73.6%; Pred. No. 2,5e-128;
Matches 721; Conservative 0; Mismatches 6; Indels 253; Gaps 2;
QY      928 TCACAGGAGACCACTACTGCGCTCTGACACACAGCCGGAGTGGCATAGCTGGCCC 987
DB      871 TTCAGTGGGACCCACTACTGCGCTCTGACACACAGCCGGAGTGGCATAGCTGGCCC 930
QY      988 ATTGCTCATAGGCGCCCAAGGCTCTTACAGAGTGGATGCTCTTCTCGGAAAGAA 1047
DB      931 ATTGCTCATAGGCGCCCAAGGCTCTTACAGAGTGGATGCTCTTCTCGGAAAGAA 990
QY      1048 AAACCTATCTGCTCAGAGTGTATTGGGGAGAGAGGCTTAGAGTGAAGACTGGACAAG 1107
DB      991 AAACCTATCTG----- 1003
QY      1108 CATATCAACTGTATTATTATTAACATCTTGTCTCCAGGACCCAGATATATGCT 1167
DB      1004 -----TCCAGGACCCAGAGTATATGCT 1027
QY      1168 TCCTGCAAGAGGAGGCTATACCTAGTAAAGGCTTATCCGAGCGGCTGGAAGAGAG 1227
DB      1028 TCCTGCAAGAGGAGGCTATACCTAGTAAAGGCTTATCCGAGCGGCTGGAAGAGAG 1087
QY      1228 TCGGAGACCCCTCATGAGGATTTATCTGAGACTCTGTGATGCGGCTTATCTGCGTGGGT 1287
DB      1088 TCGGAGACCCCTCATGAGGATTTATCTGAGACTCTGTGATGCGGCTTATCTGCGTGGGT 1147
QY      1288 CTCTGCGCTCATATCATATGACAGTGAAGGAGGCTTGTGGTCTTGAAGGAGCACTTTGTT 1347
DB      1148 CTCTGCGCTCATATCATATG----- 1168
QY      1348 CTCTACCTGTCTGTGTGGCATAGATCCCAACAGGAGCATGAGAGGCTTAGTGAAGATCC 1407
DB      1169 ----- 1168
QY      1408 CCAAGGACATGAGAGGCTTAGTCAAGATCCCATGACATGAGAGCATGTATTGG 1467
DB      1169 ----- 1168
QY      1468 TGCCTTCTCCAGAGCGGCGTGTGTGCTGACCTGAACTGAGAGCCCAAGCCAC 1527
DB      1169 -----CAGACGCGGCGTGTGTGCTGACCTGAACTGAGAGCCCAAGCCAC 1217
QY      1528 GTGAGACAGAGCTTCTTGGGCCCATGAGAAAGTGAACGAGCCTTGTGTATGAAAAAGTC 1587
DB      1218 GTGAGACAGAGCTTCTTGGGCCCATGAGAAAGTGAACGAGCCTTGTGTATGAAAAAGTC 1277
QY      1588 CTTGGCCCTTAATCATATGTTCCGCAATGATGCCGCTTGTACTCATATCATATGATCCAA 1647
DB      1278 CTTGGCCCTTAATCATATGTTCCGCAATGATGCCGCTTGTACTCATATCATATGATCCAA 1337
QY      1648 TTTGTACTGCTACATGATATGAGAAATGAAATGAGCAAGGCCCTTCCGCAACCCA 1707
DB      1338 TTTGTACTGCTACATGATATGAGAAATGAAATGAGCAAGGCCCTTCCGCAACCCA 1397
QY      1708 GAATGTGACAGCTCTCTGTGGCTGACCTACATGAGAGGCGCTTGTGACATGAGCTGGCCT 1767
DB      1398 GAATGTGACAGCTCTCTGTGGCTGACCTACATGAGAGGCGCTTGTGACATGAGCTGGCCT 1457
QY      1768 GGGCCACCTCTAGTTCCTCATATTAAGACAGATTGCTTTCCTTCACTGAGGG 1827

```


QY 1038 CTGGAGAGAAAACTATCTGGCCAGGTGTGATTTGGGGAGAGGCTTGGATGAGA 1097
 DB 173736 CTGGAGATGAGAAAGCTATCTGATCCAGGTTTGTATTTGGAGAAAGGCTTAAGATGAAA 173795
 QY 1098 CTGGAGCAAGACATATCCACTCTGTATTTATTAACATCTTTGCTCTCCAGGACCCAG 1157
 DB 173796 TAGGAGCATGGCTAT--GGGTATATGTATTAATCTCTTTGTTCCCTTACGGGCACTCA 173853
 QY 1158 GTATATGCTCTTCTGACAAAGGAGGTATACCTTAAGACGGTTATCCGAAGCGGCTG 1217
 DB 173854 GTATATGCTCTTCTGACAAAGGAGGTATACCTTAAGATGATGATTAATCAAGACGGCTG 173913
 QY 1218 GAGAGAGAGTGGGAGACCCCTCAATGGGATATCTGTGACCTGTGATGAGTGCGGCTTATC 1277
 DB 173914 GAGAGAGATTTGGAGACCTTCCCGGATCAAGCTTGTATACATAGATGAGGCTTTTCC 173973
 QY 1278 TGGCTGGGTCTTCTCGGCTCCATATCATGAGAGGAGGCTTGTGAGCTTATGAGG 1337
 DB 173974 TGGCTGGGTCTTCTCGAGCTCTAGCATCATGAGTGAAGGACCTTTGGGTTAGAGG-GG 174032
 QY 1338 GCACTTTGTTCTGCTACTCTGTGAGCATGATCCCAACAGGAGCATGAGAGGCTTAG 1397
 DB 174033 GTGCTTTGTTCTTCTTAATCTTCAAGCATAGAT----- 174066
 QY 1398 GTGAGATCCCGGAGCATGAGAGGCTAGGTAGGATCCCATGACATGAGAGCATG 1457
 DB 174067 -----GATCAGAGAGGAGCATGAGAGAGCTATATTCAGATTTGACCGATGAGAGGCTATG 174121
 QY 1458 CTATGTTGTTGCTCTTCTCCCAAGAGCGGCGCTGTGAGCTGAGACCTGATGAGAG 1517
 DB 174122 TTAAAGTTGTTGCTCTTCTCCAGAGCGGCGGCTTGTGAGCTGAGACCTGATGAGAG 174181
 QY 1518 CCCAAGCCAGTGTGACAGAGCTCTTGTGCGCCATGAGAAAGTATGACGAGCCTTGTA 1577
 DB 174182 CCCAAGCCAGTGTGACAGAGCTCTTGTGCGCCATGAGAAAGTATGAGCCTTGTA 174241
 QY 1578 TGGAAAGTCCCTGGCCCTTACTCATGTTCCGCAATGGTCCCGGCTTACTCATTC 1637
 DB 174242 TGGAAAGTCCCTGGCCCTTACTCATGTTCCGCAATGGTCCCGGCTTACTCATTC 174301
 QY 1638 ATGTGTCATATTTGATCTGCTAGACGATGTGAGAAAGTAAATGACGAGCCCTTC 1697
 DB 174302 ATGTGTCATATTTGATCTGCTAGACGATGTGAGAAAGTAAATGACGAGCCCTTC 174361
 QY 1698 CGAAGCCCAAGATGTGACAGAGTCTCTGCGGCTGCACTCACTGAGGGGCC---TTCTGAC 1754
 DB 174362 CTCAGGCCCGAGAAAGTGAACAGCATCTTGGCTGCACTCACTCAATAAAAAGCCTGATGGGAA 174421
 QY 1755 ATGAGTCTGGCCCTGGGCCCACTCTAGTTCCTCATTAATAAGACGATTTGCTTCCG 1814
 DB 174422 TTAGCCAGAGCCAGCCCACTTCTCATTTCTAATAAACAAGATGTTTCTTCA 174481
 QY 1815 TT 1816
 DB 174482 AT 174483

RESULT 9
 AC131626 275982 bp DNA linear HTG 10-MAY-2003
 LOCUS AC131626/c
 DEFINITION Rattus norvegicus clone CH230-163P16, *** SEQUENCING IN PROGRESS
 *** 15 unmerged pieces.
 AC131626
 AC131626.5 GI:30520421
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 275982)
 Muzny,D,Marle,M, Metzker,M, Lee,A, Abramson,S, Adams,C, Alder,J,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anylebech,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Bunay,C., Burch,P., Butrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyler,M., Crete,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Dlyva,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Bayes,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,J.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
 Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kralt,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Louieged,H., Lozada,R.J., Lu,X., Ma,D.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Mallory,K., Mangum,A.,
 Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
 Manning,S., McLeod,M.P., McNeill,T.Z., Meenan,B.,
 Milobavlevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwokeleleh,O., Okunolu,G., Olampunsgoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C.,
 Plopper,F., Polidexter,A., Popovic,D., Pirms,B., Pu,L.,L.,
 Puzo,M., Quito,J., Rachlin,E., Reeves,K., Regier,M.A., Reich,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodrey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Saverly,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajls,D.,
 Sneed,A., Sodergren,B., Song,X.-Z., Sorelle,R., Soza,P.,
 Steidle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tinney,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G. and Gibbs,R.A.
 Direct Submission
 TITLE
 JOURNAL
 REFERENCES
 2 (bases 1 to 275982)
 Rat Genome Sequencing Consortium.
 Direct Submission.
 TITLE
 JOURNAL
 REFERENCES
 Submitted (25-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 275982)
 Rat Genome Sequencing Consortium.
 Direct Submission.
 TITLE
 JOURNAL
 REFERENCES
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 COMMENT
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angulano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismuel, K., Blair, J., Blankenburg, K., Blych, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Creg, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, T., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Herandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howell, S., Huily, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louised, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelam, O., Okunolu, G., Olarunmugbon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Pu, M., Pu, M., Quirz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soes, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, Y., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 238748)
Worley, K.C.

Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238748)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265319. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GXYV
Center clone name: CH230-25H1

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 220363 bases at least Q40
Consensus quality: 223613 bases at least Q30
Consensus quality: 226021 bases at least Q20
Estimated insert size: 223050; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 238748: contig of 238748 bp in length.
Location/Qualifiers
1. 238748
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-25H1"
1904. 2550
/note="clone_boundary
clone_end:77
site:
end_sequence: BH303560"
43040. 44214
/note="wgs contig"
complement(124965. 236969)
/note="clone_boundary
clone_end:5p6
site:
end_sequence: BH303561"
142512. 143632
/note="wgs and extension
clone_end:5p6"

misc_feature
BASE COUNT 64699 a 51467 c 52376 g 59689 t 10517 others
ORIGIN

Query Match 13.9%; Score 520; DB 2; Length 238748;
Best Local Similarity 73.3%; Pred. No. 7.6e-152;
Matches 749; Conservative 0; Mismatches 235; Indels 38; Gaps 5;

798 CTTTCACCTGGCCTTTCATCTTGGCCTTGATGATCTTCTCTCATGACTGA 857
173497 CTTTTCACCTGACCTTCTTCTTGGCCCCCGATACATCATCTCTTATGACTGA 173556
858 TTTCGATTCATCATCTACCTCTTCTCTGCGCTTCTGCGGCTTCTGCGAGGCCCTTGAAGCAA 917
173557 AACCTACCTCCATCTTCTGCTTCTTACATGAAAGTCTTACACCACTGAGCACTA 173615
918 CCTATGATTCACAGGACCACTACCTGAGGCTGACACACAGCGGAGATGCTGAGCA 977
173616 CCTGATATCCACAGGCTCCACATGAGGCTGAGCTCCAGCCGATGAGGAGCA 173675
978 TAGCTGGCCCATTTGCTCATCACTGAGGCCAGGAGTCTTCAAGAGTGAATGCTGCTTTTC 1037
173676 TAGCTGGCCCATTTGCTCATCACTGAGGCCAGGAGTCTTCAAGAGTGAATGCTGCTTTTC 173735

QY 1756 TGAGTGGCTGGCCCCCACTCTAGTTCCTCATATAAAGACAGATTGCTTGGCT 1815
 DB 131078 TGATATGGCCAGCCACCCTCTCATTTTCATCTAATAAGCAATGTTCTTCACA 131137
 QY 1816 T 1816
 DB 131138 T 131138

RESULT 7
 MMEPEX8 1467 bp DNA linear ROD 29-JUL-2002
 LOCUS M.musculus gene for hemopexin (exons 8-10).
 DEFINITION X56830
 ACCESSION X56830.1 GI:22022647
 VERSION haem binding; hemopexin; plasma glycoprotein.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 372 to 1467)
 Nucleotide sequence of the mouse haemopexin gene
 2 (bases 1 to 1467)
 Alam, J.
 Direct Submission
 Submitted (10-DEC-1990) Alam J., Alton Ochsmar Medical Foundation,
 Department of Molecular Genetics, 1516 Jefferson Highway, New
 Orleans, LA 70121, USA
 Location/Qualifiers
 1..1467
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1467
 /number=9
 intron join(461..591,675..837,975..1234,461..591,675..837,
 mRNA prim_transcript <1..>1467
 975..>1234)
 /usedin=X56829:mRNA_hmpex
 461..591
 /number=8
 exon 592..674
 /number=7
 intron 675..837
 /number=9
 exon 838..974
 /number=8
 intron 975..>1234
 /number=10
 exon 359 a 364 c 349 g 395 c
 BASE COUNT
 ORIGIN

Query Match 14.9%; Score 557.8; DB 10; Length 1467;
 Best Local Similarity 73.8%; Pred. No. 3e-164; Indels 40; Gaps 2;
 Matches 754; Conservative 0; Mismatches 227;

QY 796 TCGTTCTCACTGGCTCTTCATCTTGAGTCATCTCTCCCTCATGACT 855
 DB 322 TTCCACTCATTTACTCTTCTTCGGGCTCTGATGATTCATCCCTCATGACT 381
 QY 856 CATTTCGATTCATCACTAGCCTCTTCTGCTGGCTTTCGCGAGGCCCTAGAGC 915
 DB 382 CAACCTGATCATCTGTCCTGCTTCACTAGAAAGCCTTACCAACAGCACTACAC 441
 QY 916 AACCTATGATTTCCACAGGAGCCCACTAGCGCTGAGACACAGCGGGATGAGTGG 975
 DB 442 AACCTTATCACTCCACAGGCTTCACCTGAGCTGAGCTCCAGCGGTATGAGTGG 501
 QY 976 CATAGTGGCCCATGCTCATAGTGGCCCAAGGTCCTTCAGCAGTGGATGCTGCTTT 1035

DB 502 CATAGTGGCCCATGCTCATAGTGGCCCTCAAGGTCCTTCAACAGTAGATGCTGCTTT 561
 QY 1036 TCGTGGAGAAATACTATCTGCTCAGAGTGTATGGGGAGAGGCTTAGAGTGA 1095
 DB 562 TCGTGGAGTAAAGTCTATCTGATCAGATGTATGGAGGAAAGCTTAAGATGA 621
 QY 1096 GACTGGACAGCATATCCAACTGTATTTATACATCTTTTCTTCAGAGGACCC 1155
 DB 622 AACAGGACAGGCTATACAGCTATGTATTAATCTCTCTTTTCCCGAGGACATC 681
 QY 1156 AGGTATATGTCTTCTGACAAAGGAGCTATACCTAGTAAGCGGTATCCGAGCGC 1215
 DB 682 AAGTATATGTCTTCTGACAAAGGAGGAGCAATPAACCTAGTAAAGTGTATCCAAAGCGC 741
 QY 1216 TGAGAGAGAAAGTCGGGAGCCCTCAGAGGATTAATCTGAGACTGTGAGATGGGCTTTA 1275
 DB 742 TGAGAGAGAAAGTCGGGAGCCCTCAGAGGATTAATCTGAGACTGTGAGATGGGCTTTT 801
 QY 1276 TCTGCTGGGCTTCTTCTCGGCTTCATATGAGGAGGAGGAGGCTTCTGAGTCTTA 1335
 DB 802 CCGGCTGCTGCTTCTTCCAGGCTCTATGCTCATAGTGAAGTCTTTGGGTACTAAG 861
 QY 1336 GGGCAGCTTGTCTGCTACCTGTCTGTGAGCATAGATCCGACCAAGGAGGCTT 1395
 DB 862 GATGCTGTGCTTTTAACTTTTCAAGGATA----- 892
 QY 1396 AGGTACAGATCCCGAGGAGATAGAAAGGCTTAGTGAATCCCAATGACAGGCA 1455
 DB 893 ----GATGACAGAGAGGAGATAGAAAGCTTAATCAAGTCTGCAATGATAGAAACTA 948
 QY 1456 TGTATGTTGGTGCCTTCTCCCAAGACGCGGCTGTGTGAGCTGAGACTGACAG 1515
 DB 949 TATTAAGTTGTGTCTTCTTCCAGAGGCGGCTTTGTGTGCTGACTGAAGTCAAG 1008
 QY 1516 AGCCCAAGCCAGCTGACAGAGCTTCTTGGCCCAATGAAAGTACGAGAGCTTGTG 1575
 DB 1009 AGCCCAAGCCAGCTGACAGAGCTTCTTGGCCCAATGAAAGTACGAGAGCTTGTG 1068
 QY 1576 TATGAAAGTCCCTTGGGCTTACATCATGTTCCGCAATGTCTCCGGCTTGACTCAT 1635
 DB 1069 TTTGACAGATCTTGTGGCCCAACAGATTTTCCAAATGATTCAGCTTGTACTTAT 1128
 QY 1636 CCATGCTCCCAATTTGATCTGCTACAGTATGTGAGAAATGAAATGACAGGACCTT 1695
 DB 1129 CCATGCGCCCAATTTGATCTGCTATGACAGTATGACAAATGATGACAGGAGTCT 1188
 QY 1696 TCCGCAACCCAGAAATGATGACAGATCTCTGAGCTGACCTAGAGGGCTTTCACA 1755
 DB 1189 GCCTGAGCTCAGAAAGTGAACAGATCTTGGCTGCAATGATGAAGAGC-----CC 1241
 QY 1756 TGAGTGGCTGGGCCCCCACTCTAGTTCCTCATATAAAGACAGATGCTTTCGCT 1815
 DB 1242 TGATATGGCCAGCCCACTCTTCATTTTCATCTAATAAAGCAGATGTTTTCACA 1301
 QY 1816 T 1816
 DB 1302 T 1302

RESULT 8
 AC120746 238748 bp DNA linear HTG 19-NOV-2002
 LOCUS Rattus norvegicus clone CH230-25H1, WORKING DRAFT SEQUENCE.
 DEFINITION AC120746
 AC120746.4 GI:25086001
 VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS _FULLTOP.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 238748)
 REFERENCE Muzny,D,Marie., Metzker,M, Lee., Abrazon,S., Adams,C., Alder,J.,

5211	7669	contig of 2459 bp in length
7670	7769	gap of unknown length
7770	13621	contig of 5852 bp in length
13622	13721	gap of unknown length
13722	19656	contig of 5935 bp in length
19657	19756	gap of unknown length
19757	31061	contig of 11305 bp in length
31062	31161	gap of unknown length
31162	40106	contig of 8945 bp in length
40107	40206	gap of unknown length
40207	50613	contig of 10407 bp in length
50614	50713	gap of unknown length
50714	61952	contig of 11239 bp in length
61953	62052	gap of unknown length
62053	74978	contig of 12926 bp in length
74979	75078	gap of unknown length
75079	92714	contig of 17636 bp in length
92715	92814	gap of unknown length
92815	124084	contig of 31270 bp in length
124085	124184	gap of unknown length
124185	151377	contig of 27193 bp in length
151378	151477	gap of unknown length
151478	178112	contig of 26635 bp in length
178113	178212	gap of unknown length
178213	208382	contig of 30170 bp in length
208383	245948	gap of unknown length
245949	249548	contig of 40966 bp in length
249549	249622	gap of unknown length
249623	249722	contig of 74 bp in length
249723	249762	gap of unknown length
249763	249766	contig of 47 bp in length.

FEATURES
SOURCE

```
misc_feature 1. .1085 /note="assembly_name:Contig33"  
misc_feature 1186. .3097 /note="assembly_name:Contig84"  
misc_feature 3198. .5110 /note="assembly_name:Contig85"  
misc_feature 5211. .7669 /note="assembly_name:Contig86"  
misc_feature 7770. .13621 /note="assembly_name:Contig87"  
misc_feature 13722. .19656 /note="assembly_name:Contig88"  
misc_feature 19757. .31061 /note="assembly_name:Contig89"  
misc_feature 31162. .40106 /note="assembly_name:Contig90"  
misc_feature 40207. .50613 /note="assembly_name:Contig91"  
misc_feature 50714. .61952 /note="assembly_name:Contig92"  
misc_feature 62053. .74978 /note="assembly_name:Contig93"  
misc_feature 75079. .92714 /note="assembly_name:Contig94"  
misc_feature 92815. .124084 /note="assembly_name:Contig95"  
misc_feature 124185. .151377 /note="assembly_name:Contig96"  
misc_feature 151478. .176112 /note="assembly_name:Contig97"  
misc_feature 178213. .208382 /note="assembly_name:Contig98"  
misc_feature 208483. .249448 /note="assembly_name:Contig99"  
misc_feature 249549. .249622 /note="assembly_name:Contig99"
```

```
misc_feature      /note="assembly_name:Contig24"
                  249723..249769
                  /note="assembly_name:Contig56"
BASE COUNT      67104 a 56647 c 55906 g 68301 t 1811 other%v
ORIGIN
```

Query Match	15.2%	Score 567.4;	DB 2;	Length 249769;
Best Local Similarity	75.0%;	Pred. No. 7.5e-167;		
Matches 766;	Conservative 0;	Mismatches 216;	Indels 39;	Gaps 3;

Qy	796	TCCTTTCACCTGGGCTCTTCCATCTTTGGCCCTGGAATGCAATCTCTCCCTCATGACCT	855
Db	130157	TTCCACATCTTCTACCTCTTCTCTTGAGGCTCTGATGCAATCTCCCTCATGACCT	130216
Qy	856	CATTCTGCAATTCACCTAGCCCTTCTCTGACCCTGCTTGCCAGCGGCTTAGC	915
Db	130217	CAACCTGCAATTCATGTCGCGCTTCACTACGAAGGCTTTACCAACAGCACTGAC	130276
Qy	916	AACTTAATGATATTCACACAGGAAACCACTACCTGAGCTCTGGAACAACGAGCTGGATGGCTGG	975
Db	130277	AACTCTTAACATCCCAACAGGCTCCCACTACGAGCCTCTGGAATCCAGCCGTATGATGG	130336
Qy	976	CATAGCTGGGCCCATTTGCTCATCAGTGGGCCCAAGGCTCCTTCAGCAAGTGAATGTGCTTT	1035
Db	130337	CATATCTGGCCCATTTGCTCATCAGCTGAGGCTCAGGGTCTTCAACAGTAATGTGCTTT	130396
Qy	1036	TCCTGGGAAGAAAACTCTATCTGTGTCAGAGTGATGATTTGGGGAGAGGCTTGAGGTAGA	1095
Db	130397	TCCTGGGATGATTAAGTCTATCTGATCCAGGATGATATTTGGAGAAAGGCTTAAGATGA	130456
Qy	1096	GACTGGGACAAAGCATATCCAACTGTGATTTATTAACATCCTTTGTCTCCAGGGCAACC	1155
Db	130457	AACAGGGCAAAAGGCTATCAGAGTCAATGTTGATTAATCTCCTTTGTTCCCAAGGGCACTC	130516
Qy	1156	AGGATATATGCTTCTCTGCAAAAGGAGGCTATACCTCTGTAAGCGGTTATCCGAAGCGGC	1215
Db	130517	AAGGTATATGCTTCTCTGCAAAAGGAGGCAATACCTGTAAGTGGTTATCCAAAGCGGC	130576
Qy	1216	TGAGAAAGGAAGTCGGGACCCCTCATGAGGATTAATCTGGAATCTGTGATGCGGCTTTA	1275
Db	130577	TGAGAAAGGAAGCTTGGAGCCCTCTGAGGATCAGCCTTGAGACATTAATGACGCTTTT	130636
Qy	1276	TCGACCCTGGGCTCTTCGAGCTCCATATCATAGCAGGTGAGGGGCTTCTGGTGTCTTGA	1335
Db	130637	CCTGCCCTGGTCTTCCAGGCTCTATGCTCATCAGGTGAGGTGCTTTTGGGTACTAAG-	130695
Qy	1336	GGGACAGTCTTCTTCTGCTACCTGTCTGTGGCATTAATCCCAACAGGCATAGAAAGGCTT	1395
Db	130696	GGATGCTCTTGTCTTTTAACCTTTTCAACGGCATTAATGACACAGAAAGGCATTAAGAAAGCT	130755
Qy	1396	AGGTCAAGATCCCAAGGGCATGAGAAGGCTAAGTCAAGGATCCCATGATAGGAAGCCA	1455
Db	130756	AATTCAAGGTG-----CATGATATGGAACCTA	130784
Qy	1456	TGCTATGTTTGGTCCCTTCTCCAGAGCGGCGCTGTGTGGCTGACCTGAAGTCAG	1515
Db	130785	TATTAAGTTTGTGCTTCTTCCAGAGCGGCGCTTGTGTGGCTGACCTGAAGTCAG	130844
Qy	1516	AGCCCAAGCCACGTGAGACAGAGCTTCTTGGGCCCATAGAAAGTATGACAGAGGCTTTTG	1575
Db	130845	AGCCCAAGCCACATGAGACAGAGGTTTCTTGGGCCCATAGAAAGTATGATGGGGCCCTTG	130904
Qy	1576	TATGAAAGAGCCCTTGGCCCTTAATCTATGTTCCGCAATGGTCCCGGCTTGTACTCTAT	1635
Db	130905	TTTGAACAAGTCTTGGCCCCCAACATGTTCTTCAATGGTTCAGCTTGTACTCTTAT	130964
Qy	1636	CCATGATCCCAATTTGTAATGCTACAGTGAATGTGAGAACTGAATGCAAGCCAGGCCCT	1695
Db	130965	CCATGAGGCCCAATTTGTAATGCTATAGAGATATAGCAAACTGAATGCAAGCCAGAGTCT	131024
Qy	1696	TCGCGAACCACGAATGATGACAGTCTCTTGGCTGCACTCACTAGGGGCTTCTTGACA	1755
Db	131025	GCTTCAGCTTCAGAAAGTGAACACATCTCTTGGCTGCAAGTCAAGTAAAGGC-----CC	131077

REFERENCE	1 (bases 1 to 577)
AUTHORS	Almudena, F., Pol, V., Restagno, G. and Silengo, L.
TITLE	Structure of the human hemopexin gene and evidence for intron-mediated evolution
JOURNAL	J. Mol. Evol. 27 (2), 102-108 (1988)
MEDLINE	88316972
PUBMED	2842511
COMMENT	Original
FEATURES	<p>source text: Homo sapiens DNA.</p> <p>Location/Qualifiers</p>
source	<p>1..577</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="genomic DNA"</p> <p>/db_xref="taxon:9606"</p> <p>/map="11p15.5-p15.4"</p> <p><1..117</p> <p>/gene="HPX"</p> <p>/note="G00-120-054"</p> <p>/number=7</p> <p>118..248</p> <p>/gene="HPX"</p> <p>/note="G00-120-054"</p> <p>/number=8</p> <p>249..331</p> <p>/gene="HPX"</p> <p>/note="G00-120-054"</p> <p>/number=8</p> <p>332..494</p> <p>/gene="HPX"</p> <p>/note="G00-120-054"</p> <p>/number=9</p> <p>495..5577</p> <p>/gene="HPX"</p> <p>/note="G00-120-054"</p> <p>/number=9</p>
exon	
intron	
BASE COUNT	111 a 156 g 150 t
ORIGIN	About 1.5 kb after segment 5.
Query Match	15.4%; Score 573.8; DB 9; Length 577;
Best Local Similarity	99.7%; Pred. No. 2,2e-169;
Matches	575; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db	818 ATCTTGAGCTCTGAGATGATTCCTCTCCCTTCATGATCTATTCGATTCATCATCAGC 877
Qy	1 ATATTGGGCTCTGAGATGATTCCTCTCCCTTCATGATCTATTCGATTCATCATCAGC 60
Db	878 CTCTTCTCTGCTGGGCTTCTGCGAGGGGCGCCCTAGAGAACTATGATTCATTCAGAGGA 937
Qy	61 CTCTTCTCTGCTGGGCTTCTGCGAGGGGCGCCCTAGAGAACTATGATTCATTCAGAGGA 120
Db	938 CCCACTACTGCGCTCTGAGACACAGCCGGATGCGTGTGACATGATGAGCCATTGCTCATC 997
Qy	121 CCCACTACTGCGCTCTGAGACACAGCCGGATGCGTGTGACATGATGAGCCATTGCTCATC 180
Db	998 AGTGGCCCCAGGGTCTTCAAGATGATGCTGCTTTTCTGTGGAGAAAACTCTATC 1057
Qy	181 AGTGGCCCCAGGGTCTTCAAGATGATGCTGCTTTTCTGTGGAGAAAACTCTATC 240
Db	1058 TGGTCCAGATGATTTGGGGGAGAGGCTTGAAGTGAAGACTGGGACAAGCATATTCAC 1117
Qy	241 TGGTCCAGATGATTTGGGGGAGAGGCTTGAAGTGAAGACTGGGACAAGCATATTCAC 300
Db	1118 TCTGATTTATTAACCATCTTCTGCTCCAGGGCACCAAGTATATGCTTCTGTGACAA 1177
Qy	301 TCTGATTTATTAACCATCTTCTGCTCCAGGGCACCAAGTATATGCTTCTGTGACAA 360
Db	1178 GGGAGGCTATACCTATGTAAGCGGTTATCCGAAGCGGCTGAGAAAGAAAGTCGGAGCCC 1233
Qy	361 GGGAGGCTATACCTATGTAAGCGGTTATCCGAAGCGGCTGAGAAAGAAAGTCGGAGCCC 420
Db	1238 TCATGGGATTAATCCGGACTGATGAGATGCGGCTTATATTCGCGCTGGGATCTTCGGGCT 1297
Qy	421 TCATGGGATTAATCCGGACTGATGAGATGCGGCTTATATTCGCGCTGGGATCTTCGGGCT 480

0y	1298	CCATATCATGGCAGGTAGAGGGGCTTCTGGGGCTTAAGAGGCACTTGTCTGCTACTG	1357
Db	481	CCATATCATGGCAGGTAGAGGGGCTTCTGGGGCTTAAGAGGCACTTGTCTGCTACTG	540
0y	1358	TCTGTGGCATAGATCCACCAGGCGCATGAGAAAGCC	1394
Db	541	TCTGTGGCATAGATCCACCAGGCGCATGAGAAAGCC	577
RESULT 6			
AC125227		249769 bp	DNA
AC125227		Mus musculus chromosome UNK clone RP23-42113,	WORKING DRAFT
DEFINITION		SEQUENCE, 19 unordered pieces.	
AC125227		AC125227.3	GI:22476035
VERSION		HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.	
KEYWORDS		HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE		1 (bases 1 to 249769)	
JOURNAL		McPherson,J.D. and Waterston,R.H.	
REFERENCE		2 (bases 1 to 249769)	
AUTHORS		McPherson,J.D. and Waterston,R.H.	
TITLE		Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park	
JOURNAL		Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park	
REFERENCE		3 (bases 1 to 249769)	
AUTHORS		McPherson,J.D. and Waterston,R.H.	
TITLE		Direct Submission	
JOURNAL		Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park	
COMMENT		Parkway, St. Louis, MO 63108, USA	
		On Aug 25, 2002 this sequence version replaced gi:22138702.	
		----- Genome Center -----	
		Center: Washington University Genome Sequencing Center	
		Center code: WUGSC	
		Web site:http://genome.wustl.edu/gsc/index.shtml	
		Contact: submissions@wustl.wustl.edu	
		Project Information	
		Center project name: M_BA0042113	
		----- Summary Statistics -----	
		Sequencing vector: M13; 0%	
		Chemistry: Dye-primers ET; 0% of reads	
		Chemistry: Dye-terminator Big Dye; 100% of reads	
		Assembly program: Phrap; version 0.990319	
		Consensus quality: 240797 bases at least Q40	
		Consensus quality: 242108 bases at least Q30	
		Consensus quality: 242698 bases at least Q20	
		Insert size: 204000; agarose-fp	
		Insert size: 248995; sum-of-contigs	
		Quality coverage: 10.08 in Q20 bases; agarose-fp	
		Quality coverage: 8.91 in Q20 bases; sum-of-contigs	
		----- NOTE: This is a 'working draft' sequence. It currently	
		* consists of 19 contigs. The true order of the pieces	
		* is not known and their order in this sequence record is	
		* arbitrary. Gaps between the contigs are represented as	
		* runs of N, but the exact sizes of the gaps are unknown.	
		* This record will be updated with the finished sequence	
		* as soon as it is available and the accession number will	
		* be preserved.	
		1 1085: contig of 1085 bp in length	
		* 1086 1185: gap of unknown length	
		* 1186 3097: contig of 1912 bp in length	
		* 3098 3197: gap of unknown length	
		* 3198 5110: contig of 1913 bp in length	
		* 5110 5210: gap of unknown length	

QY 181 AAAAAAAAAAAGATACTGGAGAGTCAAGGTGTCACAGAAAGCCAGTTCCAAAAAA 240
Db 20188 AAAAAAAAAAAGATAGAGATACCTGGAGAGTCAAGGTGTCACAGAAAGCCAGTTCCAAAAAA 20247
QY 241 GACATTTTAAAGGAAAGAAAGTGTGAGCAGTCCAGTCTCTGAGAGGTAGGTCAAT 300
Db 20248 GACATTTTAAAGGAAAGAAAGTGTGAGCAGTCCAGTCTCTGAGAGGTAGGTCAAT 20307
QY 301 GAGAACAGAGAAATTCACATGAGATTTCCGAAATTGGAGAAATCTAGCAACCTGATAG 360
Db 20308 GAGAACAGAGAAATTCACATGAGATTTCCGAAATTGGAGAAATCTAGCAACCTGATAG 20367
QY 361 AACCAATTTCAATGTTGAGGAGAAAGAAAGTGTATTAAGAGATTTAGAGAAAAAGAC 420
Db 20368 AACCAATTTCAATGTTGAGGAGAAAGAAAGTGTATTAAGAGATTTAGAGAAAAAGAC 20427
QY 421 AATGGAGCTTAGATTAATTCCTTAATTAAGTGTGTGAAAAGAGAAAGAAAAAGCGG 480
Db 20428 AATGGAGCTTAGATTAATTCCTTAATTAAGTGTGTGAAAAGAGAAAGAAAAAGCGG 20487
QY 481 GGTCTAGCCAGTACTCCCTCACTCTCCACCACTCATAGGAGAGATGAGAAAC 540
Db 20488 GGTCTAGCCAGTACTCCCTCACTCTCCACCACTCATAGGAGAGATGAGAAAC 20547
QY 541 AGCCAGAGTGAAGACATTCAGTAGAAGTGTGCTCTTTTAAGTTGAGCACTGTAT 600
Db 20548 AGCCAGAGTGAAGACATTCAGTAGAAGTGTGCTCTTTTAAGTTGAGCACTGTAT 20607
QY 601 TTCAATATCTATAACCGCATCTCTGTATCATGACACCTGAATCTTTAGGAAGTCCCG 660
Db 20608 TTCAATATCTATAACCGCATCTCTGTATCATGACACCTGAATCTTTAGGAAGTCCCG 20667
QY 661 CAACCCATGATGTTGGCTTACCTGAGAACTTAGCACTGTTTCAACCTGGCTTC 720
Db 20668 CAACCCATGATGTTGGCTTACCTGAGAACTTAGCACTGTTTCAACCTGGCTTC 20727
QY 721 TTTCAAGCACTGCTGATTCAGATTTCAAGCAAGGAGCAAGTCCCAACATGTCAGCA 780
Db 20728 TTTCAAGCACTGCTGATTCAGATTTCAAGCAAGGAGCAAGTCCCAACATGTCAGCA 20787
QY 781 GTCTTGCTCATTTCTCTTCTCACTGAGCTCTTTCATCTTGAGCTCTGAGATGCT 840
Db 20788 GTCTTGCTCATTTCTCTTCTCACTGAGCTCTTTCATCTTGAGCTCTGAGATGCT 20847
QY 841 CTCCCTCTGATGATCTGATTTCTGATTCATCACTACCTCTTCTGAGCTGCTCTG 900
Db 20848 CTCCCTCTGATGATCTGATTTCTGATTCATCACTACCTCTTCTGAGCTGCTCTG 20907
QY 901 CAGCGGCTCTAGCAACCTATGATTCACAGGAGCCACTAGCGCTGAGACAC 960
Db 20908 CAGCGGCTCTAGCAACCTATGATTCACAGGAGCCACTAGCGCTGAGACAC 20967
QY 961 AGCCGGATGCTGAGCAATGCTGAGCTCTGATCACTGAGGCTGAGGCTCTGAGCA 1020
Db 20968 AGCCGGATGCTGAGCAATGCTGAGCTCTGATCACTGAGGCTGAGGCTCTGAGCA 21027
QY 1021 GTGGAATGCTGCTTTCTCTGAGAAAGAAACTCTATCTGCTCAAGTGTGTATGGGGA 1080
Db 21028 GTGGAATGCTGCTTTCTCTGAGAAAGAAACTCTATCTGCTCAAGTGTGTATGGGGA 21087
QY 1081 GAGGCTTGAAGTGAAGTGGGACAGCATATCTCACTGTATTTATTAACATCTTTG 1140
Db 21088 GAGGCTTGAAGTGAAGTGGGACAGCATATCTCACTGTATTTATTAACATCTTTG 21147
QY 1141 TCTCTCAGGCAACCAAGTATATGTTCTCTGACAAAGGAGGCTATACCTATAGAGG 1200
Db 21148 TCTCTCAGGCAACCAAGTATATGTTCTCTGACAAAGGAGGCTATACCTATAGAGG 21207
QY 1201 GTTATCCGAAGGCTGAGAAAGAAAGTGGGACCCCTCATGGGATTTATCTGACTCTG 1260
Db 21208 GTTATCCGAAGGCTGAGAAAGAAAGTGGGACCCCTCATGGGATTTATCTGACTCTG 21267
QY 1261 TGGATGCGGCTTTATCTGCTGAGGTCTTCTGAGCTCAATATCATGGGAGTGAAGG 1320

Db 21268 TGGATGCGGCTTTATCTGCTGAGGTCTTCTGAGCTCAATATCATGGGAGTGAAGG 21327
QY 1321 TTCTGGGCTTATAGAGGAGACTGTTCTGCTACCTGTCTGTGGCTATATCCCAACAG 1380
Db 21328 TTCTGGGCTTATAGAGGAGACTGTTCTGCTACCTGTCTGTGGCTATATCCCAACAG 21387
QY 1381 GGCATGAGAAAGGCTAGTCAAGATCCCAAGGACATGAGAAAGGCTAGATCCCC 1440
Db 21388 GGCATGAGAAAGGCTAGTCAAGATCCCAAGGACATGAGAAAGGCTAGATCCCC 21447
QY 1441 ATGACATGGAAGCATGCTATGTTGGTCTCTTCTCCAGGACGCGGCTGTGAGCT 1500
Db 21448 ATGACATGGAAGCATGCTATGTTGGTCTCTTCTCCAGGACGCGGCTGTGAGCT 21507
QY 1501 GGACTGAAGTCAAGAGGCCCAAGCCAGTGAACAGAGCTCTTGCCCACTGAGAAAGT 1560
Db 21508 GGACTGAAGTCAAGAGGCCCAAGCCAGTGAACAGAGCTCTTGCCCACTGAGAAAGT 21567
QY 1561 AGACGAGGCTGTGTATGGAAGAAAGTCCCTGGCCCTTAATGATGTTCCGCCAATGCTC 1620
Db 21568 AGACGAGGCTGTGTATGGAAGAAAGTCCCTGGCCCTTAATGATGTTCCGCCAATGCTC 21627
QY 1621 CGGCTTGATCTCATCATGATGTCCTCAATTTGATCTGTAAGTATGAGAACTGAA 1680
Db 21628 CGGCTTGATCTCATCATGATGTCCTCAATTTGATCTGTAAGTATGAGAACTGAA 21687
QY 1681 TGCAGCCAGAGGCTTCCGCAACCCAGAAATGACACATCTCTGAGGCTGCACTACG 1740
Db 21688 TGCAGCCAGAGGCTTCCGCAACCCAGAAATGACACATCTCTGAGGCTGCACTACG 21747
QY 1741 AGGAGGCTTCTGACATGATGTCGCTGAGCCCACTCTGATGCTCAATTAAGACA 1800
Db 21748 AGGAGGCTTCTGACATGATGTCGCTGAGCCCACTCTGATGCTCAATTAAGACA 21807
QY 1801 GATGCTTCTGCTCTGCTCTGATGAGGAGCTTCTGACATGAT-CTGAGCTGAGCCAC 1859
Db 21808 GATGCTTCTGCTCTGCTCTGATGAGGAGCTTCTGACATGATGATGCTGAGCCAC 21867
QY 1860 TCCCACTTCTCATATATTAAGACAGATGCTTCTGATGATCAAGGACCTTGCTC 1919
Db 21868 TCCCACTTCTCATATATTAAGACAGATGCTTCTGATGATCAAGGACCTTGCTC 21927
QY 1920 GTGAAACATCTTCTTCTGATGAAAGTGAAGTGAAGTCTGCTTGAAGGCTGAG 1979
Db 21928 GTGAAACATCTTCTTCTGATGAAAGTGAAGTGAAGTCTGCTTGAAGGCTGAG 21987
QY 1980 CTCAACCAAGGCTGAGAAACAGAGGAGGAGCACTAAGGAGCAACTATCTCTG 2039
Db 21988 CTCAACCAAGGCTGAGAAACAGAGGAGGAGCACTAAGGAGCAACTATCTCTG 22047
QY 2040 CAGATGATCTTGAAGTCCAGATCATTAAGTCTTTGACAGATCTTCACTATGAG 2099
Db 22048 CAGATGATCTTGAAGTCCAGATCATTAAGTCTTTGACAGATCTTCACTATGAG 22107
QY 2100 GGGGAAAGAAACAGAGTGAAGGAGAAAGCTGAGTTTATCAAGCAAGTGAAGG 2159
Db 22108 GGGGAAAGAAACAGAGTGAAGGAGAAAGCTGAGTTTATCAAGCAAGTGAAGG 22167
QY 2160 GAGCTAGACTGCTGATTCATCTGCTGATGATGAGCAACCAAGCCAGAGCATCAG 2219
Db 22168 GAGCTAGACTGCTGATTCATCTGCTGATGATGAGCAACCAAGCCAGAGCATCAG 22227
QY 2220 AAACAAAGGCTGAGGCACTATGATGCAAGAGAGATGCTCAGAGTGTGAGG 2279
Db 22228 AAACAAAGGCTGAGGCACTATGATGCAAGAGAGATGCTCAGAGTGTGAGG 22287
QY 2280 TAGAGGCTCAATCTGAGGAGCTCAATTAAGTGTGAGGATTCAGATTAACCACT 2339
Db 22288 TAGAGGCTCAATCTGAGGAGCTCAATTAAGTGTGAGGATTCAGATTAACCACT 22347
QY 2340 CATGCTGTGTTGACATGATGATGAGCAAGTGAAGTGAAGTGGGAAAC 2399

Qy	7830	TTGATATTTTTTTGGGTTACATGTAATTTATTTATTTATTTGATTTTATTTAGTA	7889
Db	36240	TTTTGTGTTTT-----TTTTTGA	36256
Qy	7890	GAGATAGGGTTTCAACATGTTGGCAGGCTGTCTCGAATCTTGACCTCAGTGAATCCG	7949
Db	36257	GAGACGGGGTTTCAACATGTTGGCAGGCTGTCTCGAATCTTGACCTCAGTGAATCCG	36316
Qy	7950	CTGCTGCTGGCTTCCCAAGTGTGAGATTACAGGCATGAGCAACAGCTGGCCCCCTC	8009
Db	36317	CCCTCTCTGGCTTCCCAAGTGTGAGATTACAGGCATGAGCAACAGCTGGCCCCCTC	36373
Qy	8010	ATAGTTTTTATCTATCTCTTGGCTTCTTCAAACTTTGGCTTGCACGTGGAATCAT	8069
Db	36374	---AGTATCATCATTTCTAGAGAGAGCTAACAGAGCTTGTGCAAGTCACTGACAGAGA	36430
Qy	8070	GTTCCTCTCCATTTTCTTCACTTCACTTATGATCTTTCAGTCTGATTTCCAACTGATACCTCC	8129
Db	36431	GGGTGAAAGAGGGGCTTGAATCTTGGGTCTT-----GCCAGGCTGTCTTCTGGAAGC	36485
Qy	8130	CTCAGTGTCTCTTTTTTCTTAGTAAGATTTTCCAGAGAGGAATGTGAATGGCCCAATCCA	8189
Db	36486	TGAAGGCTCTCTTTTGTGTAATTCATTCAGCCAGAGATGAGAGCAGTGTGGCCCTGTGTG	36545
Qy	8190	TATTTTCAGACCAACCACTTAAAGTGTGATTTGCCAGCTATGTAATTTGGCTTACATTA	8249
Db	36546	AGGGCAGCTGTCCCAACCTTGGCAGAGGACCAAGTGTGCCATCTTGTGCAAGAGTACCA	36605
Qy	8250	ATGGGTTTGGGAATCATCATTTTACTTCTTCACTTSCACAAAGCAGCATAGCTGTGTTCTCAA	8309
Db	36606	CACCTCTCTGATGTCTCTATCTTGAATGCTCAGTGAATGCTCCAGGCCCAAGAACTCAA	36665
Qy	8310	ATAGGCGCCCTTGCGCCAGGTGTGTGCTCAATGCTTAAATCCCAACATGTGTGAGGCC	8369
Db	36666	ACCTCAGCTGT--GGCCGGGCGTGTGTGCTCACACTGTAAATCCAGACCTTGTGAGGCC	36724
Qy	8370	GAGGGGGGCAATCACTTGAATCCAGAGTTCTTGAACAGCTTGGGCAACATGTGTGAAT	8429
Db	36725	GAGCGGGGTGTATCACTTGAAGTGTGAGACCAAGCTGTGCCCAACCTGTGTGAAC	36784
Qy	8430	CTCATCTCTAATAAAATACAAAAATTAGCCAGGTGTGTGTCATGACACAGTAGTCCC	8489
Db	36785	CCCGTCTCTAATAAAATACAAAAATTAGCAAGCGGTGTGTGTCATGACACAGTGTATCCC	36844
Qy	8490	AGCTGTTCAAGAGGCTGAGGTGTGAGGATTTCTCGATGTGTGAGGAGAGATTTGCAGTG	8549
Db	36845	AGTACTCTGGAAGGCTGAGGAGGATTTGCTTGA--ACCAAGAGGGGAGGTTTGAAGTG	36903
Qy	8550	AAACGTGACTGTGCTCTGCAATCAAGCTTGGGTGAC--AGATTGAGACCTGTGTCAA	8608
Db	36904	AGCGAGATGTGTCTACCTGCAATTCAGCTGTGGCAACAGAGTGAATCTCACTCAAAA	36963
Qy	8609	AAACAATTAATAAATAAA 8627	
Db	36964	AAACAAACAAACAAAAA 36982	
RESULT 8			
US-09-539-333D-1			
Sequence 1, Application US/09539333D			
Patent No. 6476208			
GENERAL INFORMATION:			
APPLICANT: Blumenfeld, Daniel			
APPLICANT: Cohen, Daniel			
APPLICANT: Chumakov, Ilya			
APPLICANT: Bouguetelerc, Lydie			
APPLICANT: Bihain, Bernard			
APPLICANT: Essioux, Laurent			
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLERIC MARKERS			
FILE REFERENCE: GNSSET.047AUS			
CURRENT APPLICATION NUMBER: US/09/539,333D			
CURRENT FILING DATE: 2000-03-30			

NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216836..216915
OTHER INFORMATION: exon V g35030 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 213818..215818
OTHER INFORMATION: 3'regulatory region g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215941
OTHER INFORMATION: exon R complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215975
OTHER INFORMATION: exon Rbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216651..216952
OTHER INFORMATION: exon Qbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216651..217061
OTHER INFORMATION: exon Q complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 217027..217061
OTHER INFORMATION: exon Q1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 229647..229742
OTHER INFORMATION: exon X complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 230408..230721
OTHER INFORMATION: exon P complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231272..231412
OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231787..231880
OTHER INFORMATION: exon O2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231870..231879
OTHER INFORMATION: exon O1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 234174..234321
OTHER INFORMATION: exon O complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 237406..237428
OTHER INFORMATION: exon Nbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239807
OTHER INFORMATION: exon N2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239853
OTHER INFORMATION: exon N complement g34872 gene
FEATURE:
NAME/KEY: exon

LOCATION: 240528..240569
OTHER INFORMATION: exon M1117 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon M52 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240824
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240994
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..241685
OTHER INFORMATION: exon M1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240800..240993
OTHER INFORMATION: exon M51 complement g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 241686..243685
OTHER INFORMATION: 5'regulatory region g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 290652..292652
OTHER INFORMATION: 3'regulatory region g34665 gene
FEATURE:
NAME/KEY: exon
LOCATION: 292653..292841

Query Match 3.0%; Score 305; DB 4; Length 319608;
Best Local Similarity 59.7%; Pred. No. 2.2e-61;
Matches 894; Conservative 0; Mismatches 454; Indels 149; Gaps 17;

QY 7173 AACACCTTATTTAAATAATACCCAGCGCGGTGTGAGTACCGCTTAATCTAGCA 7232
DB 249487 AAGTCTTTAAAGACTACATTTGAGACCGGGACCGGTGCTCATGCTTAATCCAGCA 249546
QY 7233 CTTTGGAGGCTGAGGCAAGTGTGCTTGAAGCGCTTGAAGCTTGAAGCTTGAAGCA 7292
DB 249547 CTTTGGAGGCGCAAGGTGGTGTGATC-----AGGAGTTGATACCA 249587
QY 7293 GCCTGGGCAACATGCG-AAAACCTCATCTTCAAAAAATATAAAATTAATGCGGTGTG 7351
DB 249588 GCCTGGCCAAATGTGTAAACCCGCTCTACTGAAATTAACAAAATTAAGCAGGTGAG 249647
QY 7352 GTATGCGTTTCTTATAGTCCATCTACTACTGAGGCTGAG-----CC 7393
DB 249648 GTTGCAAGGCGCTTAAATCCAGCTACTCAGAGGCTGAGGAGAGAAATCCCTTGAAC 249707
QY 7394 CGAAAGCTGAGGCTTCAAGTACCGGTGATCGTCTACTGCACTCCAGCTGGGTGAC-A 7452
DB 249708 TGGGAGGCGAGATTTGCAAGTGTGAGATTGCAACATTTGCACTCCAGCTGAGAAACA 249767
QY 7453 GAGTGAGCAATGTCTCAAAAAAACAACAACAACAACAACAACAACAACAACA----- 7508
DB 249768 GAGCAAGATTTGCTTCANAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 249827
QY 7509 -----CAAAACAAAAAACCATATATATATATATATATATATATATATATATATAT 7561
DB 249828 GAGAGTTTAAGCTATCAAAATTTAGAGATTAATGTTAATAATATGTTTGTATGAA 249887

[illegible]

```

1  RESULT 9
2  US-09-679-409-1
3  : Sequence 1, Application US/09679409
4  : Patent No. 655316
5  : GENERAL INFORMATION:
6  : APPLICANT: Cohen, Daniel
7  : APPLICANT: Blumenfeld, Marla
8  : APPLICANT: Chumakov, Ilya
9  : APPLICANT: Bougneloret, Lydie
10 : APPLICANT: Basioux, Laurent
11 : TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS
12 : FILE REFERENCE: 53.US15.CIP
13 : CURRENT APPLICATION NUMBER: US/09/679,409
14 : PRIOR FILING DATE: 2000-10-03
15 : PRIOR APPLICATION NUMBER: 09/539,333
16 : PRIOR FILING DATE: 2000-03-03
17 : PRIOR APPLICATION NUMBER: 09/416,384
18 : PRIOR FILING DATE: 1999-10-12
19 : PRIOR APPLICATION NUMBER: 60/168,088
20 : PRIOR FILING DATE: 1999-11-30
21 : NUMBER OF SEQ ID NOS: 134
22 : SOFTWARE: Patent.pm
23 : SEQ ID NO 1
24 : LENGTH: 319608
25 : TYPE: DNA
26 : ORGANISM: Homo sapiens
27 : FEATURE:
28 : NAME/KEY: misc.feature
29 : LOCATION: 199122..201122
30 : OTHER INFORMATION: 5' regulatory region
31 : NAME/KEY: exon
32 : LOCATION: 201123..201234
33 : OTHER INFORMATION: exon S
34 : NAME/KEY: exon
35 : LOCATION: 201123..201560
36 : OTHER INFORMATION: exon S2
37 : NAME/KEY: exon
38 : LOCATION: 214676..214793
39 : OTHER INFORMATION: exon T
40 : NAME/KEY: exon
41 : LOCATION: 215702..215746
42 : OTHER INFORMATION: exon U
43 : NAME/KEY: exon
44 : LOCATION: 216836..216994
45 : OTHER INFORMATION: exon V
46 : NAME/KEY: exon
47 : LOCATION: 216836..217077
48 : OTHER INFORMATION: exon V2
49 : NAME/KEY: exon
50 : LOCATION: 217671..217764
51 : OTHER INFORMATION: exon V1
52 : NAME/KEY: exon
53 : LOCATION: 227655..227736
54 : OTHER INFORMATION: exon V4
55 : NAME/KEY: exon
56 : LOCATION: 238715..238919
57 : OTHER INFORMATION: exon V3
58 : NAME/KEY: exon
59 : LOCATION: 240440..240673
60 : OTHER INFORMATION: exon W
61 : NAME/KEY: exon
62 : LOCATION: 240440..241153
63 : OTHER INFORMATION: exon W2
64 : NAME/KEY: exon
65 : LOCATION: 241072..241291
66 : OTHER INFORMATION: exon X
67 : NAME/KEY: exon
68 : LOCATION: 244353..244561
69 : OTHER INFORMATION: exon Y
70 : NAME/KEY: exon
71 : LOCATION: 246273..247802
72 : OTHER INFORMATION: exon Z

```

Query Match	Best Local Similarity	Score	DB 4	Length	319608
Matches	894	Conservative	0	Mismatches	454
				Indels	149
				Gaps	17
Query 7173 AACACCTATATTTAAATAACCCAGCGCGGCGTGAGTCAGCGCTTAAATCCTAGCA	249487	AAGCTTTAAAGACTACATCATTGAGACCGCGGCGAGCGTGTCTTAAATCCAGCA	249546		
Query 7233 CTTTGGAGGCTGAGGACGAGTGATGCTTGAGCGGCTTGACCTAGAGATTGAGACCA	249547	CTTTGGAGGCGCAAGTGGGATGATC-----AGGATTCGATACCA	249587		
Query 7293 GCGTGGGCAACATGCGC-AAAACCTCATCTTCAACAAAATATAAATTAATGTCGGGTG	249588	GCGTGGGCAACATGCTGTAAGAACCCCGTCTCACTGAAATATCAAAATTAAGCCAGGTGAG	249647		

OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70

FEATURE:
NAME/KEY: allele
LOCATION: 93690..93735
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Qy	8373	GGGGGCAATCACTTAGTGCAGAGATTCTAGACCAAGCTGGGCAACATGGTAAATCTC	8432
Db	36321	GAGGGCAGATCACTGTAGGTTTAGAGAGTTCCAGACCACTGGCCAACTGGTAAACCCC	36262
Qy	8433	ATCTCTACTTAAAA--TACAAAAAATTAGCCAGGTGTGTGGCATGCACAGTAGTCCCA	8490
Db	36261	GTCTCTACTTAAAAACAACAAAAAATTATGCGCGGGCGTGTGGCGGGCGCTGTATGTTCCCA	36202
Qy	8491	GCTTGTACGAGAGGCTGAGTGTGGAGAGATTCTCGAGTGTGGAGAGCAGAGATTGACGTGA	8550
Db	36201	GCTACGGGGGAGCGCTGAGGCGCAGAGAGATGTGCGTAACCCGGGAGGCGGAGCTTTCAGTGA	36144
Qy	8551	ACCGTGACTGTGCTCTGTGCATTCGACCTGGGTGACACAGATTGAGACCCTGTCTCAAAAAA	8610
Db	36141	GCCGAGATCGGCCACATGCACCTCCAGCGCTGGGCGACAGACGAGACTCCGTCTCAAAAAA	36082
Qy	8611	CAATTAATTAATAATAATA 8630	
Db	36081	AAAAAAAAAAAAAAAAAAAA 36062	

RESULT 12
 US-09-078-294-3/C
 : Sequence 3, Application US/09078294
 : Patent No. 6265211
 : GENERAL INFORMATION:
 : APPLICANT: Choo, Kong-Hong Andy
 : APPLICANT: Du Sart, Desiree
 : APPLICANT: Cancilla, Michael R.
 : TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
 : FILE REFERENCE: Davies Col
 : CURRENT APPLICATION NUMBER: US/09/078,294
 : CURRENT FILING DATE: 1998-05-13
 : NUMBER OF SEQ ID NOS: 29
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 3
 : LENGTH: 80595
 : TYPE: DNA
 : ORGANISM: Nucleotide sequence of HC-contig
 : US-09-078-294-3

	Query Match	3.0% ; Score 296 ; DB 3 ; Length 80595 ;
	Best Local Similarity	Pred. No. 1.5e-59 ;
	Matches 608 ; Conservative 0 ; Mismatches 230 ; Indels 68 ; Gaps 7 ;	
QY	7681 ATGAGAGTCTCACTCTGTTCACCCAGGCTGGAGTGCATATGGCGCAATCTTGTTCACTGC	7740
Db	39668 ATGAGAGCTTGGCTCTGTG-TGGCAGAGCTGGAGTGCAGTAGC---ATCTCGGGTATCTGC	39612
QY	7741 CCTCTGCTCTCTGGGGTTCAGAGCAATCTCTCTGCTCAGCCTCCCGTATACTCTGGAGTAC	7800
Db	39612 CCTCACCTCTCTAGGTTCAAGCAATTTCTCTCCCTCAACTCTCGAGTATGTTGGAGCTAC	39555
QY	7801 AGGCGCATATCCACCATGCTGGCTAATTTTTTGTATTTTTTTGGGTTCAATATGACTAT	7860
Db	39552 AGGTCATATGCACACCAACTGGCTAATTTTTTGGC-----	39511
QY	7861 TTATTAATTTAATTTTTGTATTTTTAGTAGATAGGCTTTCACATGTGGCCAGGCTG	7920
Db	39518 -----TTTTTAATAGAGACAGGGTTTTCACATGTTGGCCAGGCTG	39478
QY	7921 GTCTCGAAGCTCTGACCGTAGGTATCCGGCTGCTGGGCGTCCCAAGTGTAGAGTTA	7980
Db	39478 GTCTTAATCTCTAAGCTCTGAGGTATCCGGCCACCTTGGCTCCCAAGTGTGGAGTTA	39411
QY	7981 CAGGCATGAGCAACACGCGCTGGCCCTCATTA-----GGTTTTATCTATCTCTTGTCT	8035
Db	39418 CAGGCAGAGGCTATCTGGCGCTGGCCCAATTCATTAAGCTTTAATTAATTAAGTATTAAG	39355
QY	8036 TCTTCACAACTTGGCTTGCACGGGAGCAATCATGTTCTCTCCACTTTCACATCTCA	8095
Db	39358 TTTTAAATTTTATTAATATGTTGTATTAATTAATTAATTAATTAATCTTAATCTCAATTAAGAGCTAC	39299

Qy	8096	TGAATCTTCAAGTCTCAAGTCCAACTGATATACCTCCCTCAGTGTCTTTTTCCTAGTAAG	8155
Dd	39298	TACTGTGTGAATTCCTTATTTCTATTCATTCAAACACTATCTCCACACTTTTTCATAGGTTTAT	39233
Qy	8156	ATTTCGAGAGAGGAATTCGAATAGGCCAGTCATATTTTTCAGACCACCACTATTAAG	8215
Dd	39238	TCCCTCTTGAAGCCATCATATTAATTTTCACTTACATATGACAGAAAGACAGCTGAGAGAGAC	39178
Qy	8216	TGGTTAATTGCCAGCCCTATGTATTTGGCTACATTAATGGGTGGGAATCATCATTTACTT	8275
Dd	39178	TAAGTCAATTTTAAAGCCGTTCAATGATGGCAGCATTTGGATTTGACTCATCAAAAGCTTG	39115
Qy	8276	CATTGCACAAAGCAG-----CATAGCTGTGTTCTCAAAATATAGGCCCTCTGGGCC	8325
Dd	39118	TCAATATGAAGACAGATTTTAAGGCCCATGTGTGTCTTTCATATATTAAGACATTTGACC	39055
Qy	8326	AGGTGTGGTGGCTCATGCTATATATCCCAACACTGTGGAGGCGAGGGGGCAGATCAC	8385
Dd	39058	AGGGGTGTGGCTCAGCGCTGTATCTCCAGCACTTTTGGGAGGCGAAGGCTGGTGATCA-	39000
Qy	8386	TTGAGTCCAGAGTTTCTAAGCCAGCTGGGCAACATGGTGAATCTCATCTCTAATTTAAA	8445
Dd	38999	-TGGGTCCAGAGATTTCAAGCCAGCTGGGCAAGATGGTGAATCCCACTCTCACTTTAAA	38941
Qy	8446	ATTCAAAAAATTAGCCAGGTGTGGTGGACATGACACAGTATGCTCCAGCTGTTGAGAGGCT	8505
Dd	38940	ATAC-AAGAAATTTAGCCGGGTGACAGTGGCAGGCTCTGTATATCCCACTACTCAGAGGCTT	38888
Qy	8506	GAGGTGGAGGAATTGCTCGAGTGTGGAGAGCAGAGATTGCAAGTGAACCGTGAATGTGCTCT	8565
Dd	38881	GAGGCAGCAAAATTCGCTGTGAACCTGGGAGGTGAGGTTGCAATGAGCTGATGATCAAGCCA	38822
Qy	8566	CTGCAATCCAGCTCGGTGTGACAGATTTGAGACCTGTCTCAAAAAAATAATTAATTAATAA	8625
Dd	38821	CTGATCTCTGCTCGGTGTGACAGAGCAAGCTCTGTCTCAAAAAAATAATTAATTAATAA	38762
Qy	8626	AAATTA 8631	
Dd	38761	ATAAAA 38756	

RESULT 13-975B-6/C
 : Sequence 6, Application US/08370975B
 : Patent No. 5622851
 : GENERAL INFORMATION:
 : APPLICANT: Maley, Frank
 : APPLICANT: Maley, Gladys F.
 : APPLICANT: Weiner, Karen X.B.
 : TITLE OR INVENTION: Human Deoxycytidylate Deaminase Gene
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 : STREET: Clinton Square, P.O. Box 1051
 : CITY: Rochester
 : STATE: New York
 : COUNTRY: USA
 : ZIP: 14603
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/370,975B
 : FILING DATE: 10-JAN-1995
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Timian, Susan J.
 : REGISTRATION NUMBER: 34,103
 : REFERENCE/DOCKET NUMBER: 20894/80
 : TELECOMMUNICATION INFORMATION:

QY	7830	TTGATATTTTTTGGGTACAAGTACTATTTATTTAATTTTTGATTTTTGTA	7889
Db	17311	-----TAAATCTTGATTTTTAGTA	17292
QY	7890	GAGATAGGGTTTACACATGTTGGCCAGGCTGTCGAACTCTGACTCAGTGATCCG	7949
Db	17291	GAACACAGGGTTTGGCTATGTTGGCCAGGCTGTCGAACTCTGACTCAGTGATCCA	17232
QY	7950	CTGCTCTGGGCTTCCCAAGTGTAGATTACAGGCATGAGCAACACGCTGGCCCTC	8009
Db	17231	CTGCTCTGGGCTTCCCAAGTGTAGATTACAGGCATGAGCAACACGCTGGCCCTC	17172
QY	8010	ATAGTTTTTATCTAATCTCTGTCCTTTTCAACAATTGGCTGCACTGAGACCATCAT	8059
Db	17171	ACTACCTTAAAGAAATCCATCTGTTCTATGATCTCGAATGACCAATCATCTGTAACCT	17112
QY	8070	GTCTCTCCACTTCTCTGACTA-----CTTCATGATCTTTCAGTCTCACTTC	8115
Db	17111	GAATCTGTTGTCTAGCCATTAATCGCCCATCTTCTCCAGGCTTTCTTTTGGGTGT	17052
QY	8116	CACTGATACCTCTCCCTCAGTGTCTCTTTTCTTCTAGTAAGATTTCCAGAGGGAATCTG	8175
Db	17051	CTGGGGCTTCATGTACTCTTCTCTCTAATGAGGTAAATCTCTTATCCAAATCTTAGG	16992
QY	8176	AATGGCCCAATGCATATTTTACAGACACACACATTAAGTGTGATGTCAGCCATATG	8235
Db	16991	AGCAGAAATGTTTTCAGATTTTCAGATTTTAAATGTGTGACTTATTAATCAACAGGTGAGCA	16932
QY	8236	TATGGCTACATTAATGGGTGGGAATCATCTTACTTATGTCACAAAGCAGCATAG	8295
Db	16931	TCCCAATTCACAACTTTGAAATCAAGAGTCTCATGAGCATTTCTTCAAGTGTATG	16872
QY	8296	CTCTGTTCTTCAAA-ATAAGGCCCTCGGCGAGGTGTGTGTCATGCTTATATCCCA	8354
Db	16871	CTGTGTGCAAAAAGTCTTTCAGATTTTGGGCGAGGTGAGTATCATGCTGTATATCCA	16812
QY	8355	ACACTGTGGAGGGCGGAGGGGGGAGATCATCTTGATTCAGAGTTTATAGACCACTCG	8414
Db	16811	GCACTTTTGGGAGGCGCAAGTGGGGAGATCGTTTATGCTCAAGATTCAGAAACAGCTTAG	16752
QY	8415	GCAACATGTTGAAATCTCATCTCTACTTAAATAACAAAATTTAGCCAGGTGTGTGCA	8474
Db	16751	GCAACATGTTGACAGCTCATCTTCTACTGAAATTAATAAATTTAGCCAGGTGTGTGCG	16692
QY	8475	TGCACAGTATGCCAGCTGTTTCAGAGGCTGAGGTGGAGATTTGCTCGAGTGTGGAG	8534
Db	16691	TGCACCTGTATGCTTCACTGATCTTGGAGGCTGAGGTGGAGATTTGCTCGAGTGTGGAG	16632
QY	8535	GCAGAGTATGCAATGAAACCTGACTGTGCTCTGCAATCCAGCTGGGTGACAGATTGAG	8594
Db	16631	GTCAGAGCTGCAATGAGAGCTGTGTTTGCACCACTGATCTGAGTGTGGGCAACAGCAAG	16572
QY	8595	ACCTGTCTCAAAAACCAATATAA	8618
Db	16571	ACCCCATCTCAGAAAAAATAAAAAA	16548
RESULT 15			
US-09-491-356C-1/C			
Sequence 1, Application US/09491356C			
Patent No. 6566061			
GENERAL INFORMATION:			
APPLICANT: Philibert, Robert A.			
APPLICANT: Ghans, Edward I.			
APPLICANT: Delisle, Lynn			
TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XOL3			
FILE REFERENCE: 9465.6US11			
CURRENT APPLICATION NUMBER: US/09/491.356C			
CURRENT FILING DATE: 2000-01-26			
PRIOR APPLICATION NUMBER: PCT/US99/09365			
PRIOR FILING DATE: 1999-04-29			
PRIOR APPLICATION NUMBER: 60/083,465			
PRIOR FILING DATE: 1998-04-29			

```

NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 55298
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc_feature
LOCATION: (485)..(485)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (838)..(838)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (16728)..(16728)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (22750)..(22750)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (22756)..(22756)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (28519)..(28519)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (44804)..(44804)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (45002)..(45002)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (54049)..(54049)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (54226)..(54226)
OTHER INFORMATION: n is not determined
US-09-491-356C-1

Query Match      2.9%; Score 294.8; DB 4; Length 55298;
Best Local Similarity 61.3%; Pred. No. 2,5e-59;
Matches 601; Conservative 2; Mismatches 325; Indels 53; Gaps 6;

QY      7641 CTGACTCAAGATTGTGGTGTGGTGCTGTTGTGAATGAGTGCATCTGTGCAC 7700
DB      16910 CTGAGTCATGCTGAAGTTTTTTTTTTTTTTTTTTTGAAGACGAGTCTCGCTGTGCAC 1685

QY      7701 CCAGGCTGAGTGCATATGGCGCAATCTTGTTCACTGCAACTCTGCTCTGGGTTCAA 7760
DB      16850 CCAAGGCTAGAGTAGTGAAGTGGCGCACCTCCGCTTAATGCAAACTCCGCTCCCGGTTCCAC 1679

QY      7761 GCATCTCTCGCTCAAGCTCCCGGTATPACTGGAGTACAAGGCGCATACACATGCGCT 7820
DB      16790 GCATCTCTCGCTCAAGCTCCCGGTATPACTGGAGTACAAGGCGCGCTGCACACGCCC 1673

QY      7821 GGCTAATTTTGTATTTTGTGGTTACATGTAATAATTTAATTTTGTGA 7880
DB      16730 AGNTAATTT-----TTTTTGTGA 1671

QY      7881 TTTTATGATAGATAGGGTTTCAACATGTTGGCGAGGCTGCTCGAATCTCTGACTTCA 7940
DB      16713 TTTTATGATAGAGACGGGGTTTCAATCGGTATMACAGAGATGTCGATCTCTGACCTC- 16651

QY      7941 GGTGATCCGCTCGCTCCGCGCTCCCAAAGTGTAGAATTACAGGATAGAACCAACAGGCC 8000
DB      16654 -GTGATCCGCGCACCTCCGCGCTCCCAAAGTGTAGAATTACAGGATAGAACCAACAGGCC 16599

QY      8001 TGCGCCCTCATAGATTATTTATCTATTCTCTTGCTCTTCAACAATTTGGCTTGACAGTG 8060
DB      16595 CGGCGATGCTTGAGTTGAGGCTATGATTTGCTATCATGATCTTCAAGACAATGAGAGGAATT- 16533

QY      8061 GACCATCATGTTCTTCGACTTTTCTACATTAATTACATGATCTTTCAGTCTGAGTTCCAAT 8120

```

Db 16536 -CTTGAAGTCTCAGGCTCTGTTTCCATCTTAGAAAGTATGGAAATTGACAGAAATG 16478
QY 8121 GATACCTCCCTCAGTTCCTCTTTTCTTAGTAAGATTCCAGAGAGAAATCTGAATG 8180
Db 16477 ATTTCAGGAAATTCAT--TCAACATTCACAGTGGGTTTCCAGTGTCTTAGGGCCCAT 16420
QY 8181 CCCAGTCCATATTTTCAGACCACACACATTAAGTGTGATTGCCAGCCTATGTAATG 8240
Db 16419 CAGAACTAGAGAACTAGAGCAGAGGCCCTGCAAGATCCAGGGTACATAACTATTTGA 16360
QY 8241 GCTACATTAATGGTGGGAATCATCTTACTTCAATTGC--ACAAGCAGATAGCT 8297
Db 16359 CCATGACCAACAGATTCAGACCCCTTACCACTTTCAGTCCCTGCACTTCTTACAGAA 16300
QY 8298 CTGGTTCCTAAATATAGGCCCCCTGGGCCAGGTGTGAGGCTCATGCTTATATCCCAAC 8357
Db 16299 CAAGGTTAATATTCTTAAGACTCAGGCTGGGACGGGTGGCTCATGCTGTATTCACACA 16240
QY 8358 CTGTGGGAGGCCGAGGGGGCAGATCACTTGAATCCAGAGTTCTAGACCAGCCTGGGCA 8417
Db 16239 CTTTGGGAGGCCGAGACAGACGAGTCAATTGAGGTAGAGTTCAAGACCAAGCCTGACCA 16180
QY 8418 ACATGGTGAATCTCATCTCTACTAATAAATACAAAAATTAGCCAGGTGTGGCATTC 8477
Db 16179 GCATGGTCAAAACCCCATCTCTATGAAAAATAC--AAAAATTAGCCAGGTGTGGTAC 16121
QY 8478 ACCAGTAGTCCAGCTGTTCAGGAGGCTGAGGTGGAGGATTTGCTGAGTGTGGGAGGCA 8537
Db 16120 ACCGTATATCCAGCTACTAGAGGTTGAGGACGAAATCACTTGAACCCAGAGAGTG 16061
QY 8538 GAGATTGCACTGAACCGTGACTGTGCTCTGCAATCCAGCCTGGGTGACAGATTGAGACC 8597
Db 16060 GAGGTGCACTGAGCTGAGATGGGGCAGCTGCACTCCAGCTTGGGTGACAGATGAGACT 16001
QY 8598 CTGCTCAAAAAACAATATA 8618
Db 16000 CTGACTCCAAAAAATA 15980

Search completed: December 16, 2003, 05:47:11
Job time : 467.684 secs

Db 181 CCATTGCCCAACACCAGAAAAGACATGCTCCGCAATGGGAGAGAGGTGATATGAG 240
Qy 241 CATTGGCTGTAGAGAGGATGGCATTTGGCCAGGCTGCCAAGAGATCAGAGGTCCACCTT 300
Db 241 CATTGGCTGTAGAGAGGATGGCATTTGGCCAGGCTGCCAAGAGATCAGAGGTCCACCTT 300
Qy 301 GCCCAGTGAACCTTGAAGAGAGGATGATGTTCAAGACATTTTCATTCGTAAAGTCAAGA 360
Db 301 GCCCAGTGAACCTTGAAGAGAGGATGATGTTCAAGACATTTTCATTCGTAAAGTCAAGA 360
Qy 361 GAGGACATTGAGCCTGATGAGAGAGGCTGTGATGATGTTGTTCCAGAGGTTCCGGAATG 420
Db 361 GAGGACATTGAGCCTGATGAGAGAGGCTGTGATGATGTTGTTCCAGAGGTTCCGGAATG 420
Qy 421 TGTGTTTTCTGTGGAAGGAACTTGTGAGAGTAAAGAGGATCTGAGACTTTTGGTGA 480
Db 421 TGTGTTTTCTGTGGAAGGAACTTGTGAGAGTAAAGAGGATCTGAGACTTTTGGTGA 480
Qy 481 AGATTATATATGAGAGCTGAGAGGCTGTGAGAGCTGTGAGAGGATCAGAGGCTTTTCA 540
Db 481 AGATTATATATGAGAGCTGAGAGGCTGTGAGAGCTGTGAGAGGATCAGAGGCTTTTCA 540
Qy 541 GCCTTGGCTAGGAGAGAGGAGGCTGTGAACTTGTGATCCTGAGCTGAGCTGAGCTGCA 600
Db 541 GCCTTGGCTAGGAGAGAGGAGGCTGTGAACTTGTGATCCTGAGCTGAGCTGAGCTGCA 600
Qy 601 TAATTTCTTTTGTACCTCACTAGGCAATCTCAACAGAAATGGGGCAGCTTTGGAGTGG 660
Db 601 TAATTTCTTTTGTACCTCACTAGGCAATCTCAACAGAAATGGGGCAGCTTTGGAGTGG 660
Qy 661 GCCCAGAGATGATGAGAGATGAGAGGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 GCCCAGAGATGATGAGAGATGAGAGGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 721 ACACTGAGAGAGAAACAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 ACACTGAGAGAGAAACAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 781 ATGCCCTAATACCTAATACCTAATACCTAATACCTAATACCTAATACCTAATACCTAAT 840
Db 781 ATGCCCTAATACCTAATACCTAATACCTAATACCTAATACCTAATACCTAATACCTAAT 840
Qy 841 ATGTGAGAAATTTGGCTGCACTAGCACTAGCACTAGCACTAGCACTAGCACTAGCACTAG 900
Db 841 ATGTGAGAAATTTGGCTGCACTAGCACTAGCACTAGCACTAGCACTAGCACTAGCACTAG 900
Qy 901 CATTTTTCATTCACAAATATTTCTGGAACAATCTTATATGCGCAGGCACTATTTTATG 960
Db 901 CATTTTTCATTCACAAATATTTCTGGAACAATCTTATATGCGCAGGCACTATTTTATG 960
Qy 961 GAGTCAAGGATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 961 GAGTCAAGGATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Qy 1021 TCACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 TCACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Qy 1081 AAGGCTGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1081 AAGGCTGTGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy 1141 AAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1141 AAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Qy 1201 GGTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 GGTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Qy 1261 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 1261 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320

Qy 1321 ACTTCAGGACATTTTGTGAAAGCCCTTTCAGAGTCAAGGAAAGGAGAGGCTGGCAGT 1380
Db 1321 ACTTCAGGACATTTTGTGAAAGCCCTTTCAGAGTCAAGGAAAGGAGAGGCTGGCAGT 1380
Qy 1381 GGGCTTGTGACATTTGACAAAGTGAAGATATATGATGATGATGATGATGATGATGATGAT 1440
Db 1381 GGGCTTGTGACATTTGACAAAGTGAAGATATATGATGATGATGATGATGATGATGATGAT 1440
Qy 1441 AAAAGTGAAGAGCTTGGGATGCCCCAGAAAGGCCCCGATGATGATGATGATGATGATGAT 1500
Db 1441 AAAAGTGAAGAGCTTGGGATGCCCCAGAAAGGCCCCGATGATGATGATGATGATGATGAT 1500
Qy 1501 TTCTGAGATATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 TTCTGAGATATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Qy 1561 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 1561 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Qy 1621 CTTTTTCTACCTTCCCTCCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CTTTTTCTACCTTCCCTCCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Qy 1681 ACTCACTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1681 ACTCACTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Qy 1741 CGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 CGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Qy 1801 CCGGAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 CCGGAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Qy 1861 TGTATCAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Db 1861 TGTATCAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Qy 1921 TGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
Db 1921 TGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
Qy 1981 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
Db 1981 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
Qy 2041 GAGAGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Db 2041 GAGAGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Qy 2101 TAGAAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 TAGAAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Qy 2161 CAAATTTGGGCTCAGGGCTATGAGAAAGGATCGGTCCCAATGATGATGATGATGATGAT 2220
Db 2161 CAAATTTGGGCTCAGGGCTATGAGAAAGGATCGGTCCCAATGATGATGATGATGATGAT 2220
Qy 2221 TTGTCTCCCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 TTGTCTCCCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Qy 2281 CAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 2281 CAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Qy 2341 TGCATTTATACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
Db 2341 TGCATTTATACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400

Qy	2401	ACCTCCACCCCAAGAACGCTGCTCAGATGCGCTGAGACCTTTGATGTCTACCAACCCCTGGATGAC	2460
Dp	2401	ACCTCCACCCCAAGAACGCTGCTCAGATGCGCTGAGACCTTTGATGTCTACCAACCCCTGGATGAC	2460
Qy	2461	AATGGAACCAACGCTGTTTTTTTAAAGATGAGAGGAGCTAGAGGTTAGGGCGTTTAGACCTT	2520
Dp	2461	AATGGAACCAACGCTGTTTTTTTAAAGATGAGAGGAGCTAGAGGTTAGGGCGTTTAGACCTT	2520
Qy	2521	AGACTTACTCTCTTCCAAAGGGGTGCTCTGTGTGGAGAGCTTAGAATTAATCTGA	2580
Dp	2521	AGACTTACTCTCTTCCAAAGGGGTGCTCTGTGTGGAGAGCTTAGAATTAATCTGA	2580
Qy	2581	TGATATCACTGACAGCTTCTCTCAAGCTATCTCAGTAGCTCAAAAGTTTCTCACTGGGCC	2640
Dp	2581	TGATATCACTGACAGCTTCTCTCAAGCTATCTCAGTAGCTCAAAAGTTTCTCACTGGGCC	2640
Qy	2641	CCTCAGTAGAGTGGGTTTTTTCAGGGGAGTTGTGTGGAAGAGTCAAAATGGGACCGG	2700
Dp	2641	CCTCAGTAGAGTGGGTTTTTTCAGGGGAGTTGTGTGGAAGAGTCAAAATGGGACCGG	2700
Qy	2701	GAGTTAAATCTCAGAGAGATGGAAGATTTCCCAAGCCCTGTGATAGTGTGCAATTCGGTCAA	2760
Dp	2701	GAGTTAAATCTCAGAGAGATGGAAGATTTCCCAAGCCCTGTGATAGTGTGCAATTCGGTCAA	2760
Qy	2761	GGTCACAACAGTGTCTTTCTGATCAAGTACTGTGGCCAAATTCAGGGCCAGGCTGGA	2820
Dp	2761	GGTCACAACAGTGTCTTTCTGATCAAGTACTGTGGCCAAATTCAGGGCCAGGCTGGA	2820
Qy	2821	AAGGGCTGGAAATGCACACTGGGGAACCTTCCCAATAGGCGCTTGCGCATGGAGCCCATAG	2880
Dp	2821	AAGGGCTGGAAATGCACACTGGGGAACCTTCCCAATAGGCGCTTGCGCATGGAGCCCATAG	2880
Qy	2881	CAATAGGTAGCAGATTTCTTTCCTCATGTGCCCTCCTTCTGTAAAGCTTGGGCTAAG	2940
Dp	2881	CAATAGGTAGCAGATTTCTTTCCTCATGTGCCCTCCTTCTGTAAAGCTTGGGCTAAG	2940
Qy	2941	GAGTGTGCATGCGTGTGGGCGCTGGCAGGTGCAACATCCAGTGCGTGTCTTCACTCTAG	3000
Dp	2941	GAGTGTGCATGCGTGTGGGCGCTGGCAGGTGCAACATCCAGTGCGTGTCTTCACTCTAG	3000
Qy	3001	TCTTAGTTCTACACCGCTGTGTGTGAACCTCAACATGCTGGGCATTCCTTTTTCCTGGC	3060
Dp	3001	TCTTAGTTCTACACCGCTGTGTGTGAACCTCAACATGCTGGGCATTCCTTTTTCCTGGC	3060
Qy	3061	AATTGCTTCCCTTGCCCTTCCATGACCCCTGTATCAAGTCTCTTCAATAGGGCAAGCAAGT	3120
Dp	3061	AATTGCTTCCCTTGCCCTTCCATGACCCCTGTATCAAGTCTCTTCAATAGGGCAAGCAAGT	3120
Qy	3121	TGTTTCCCAACAATAGGCACTGGCTAGAGAGCAATGTGAGCAATGAATCAAGTCTGCT	3180
Dp	3121	TGTTTCCCAACAATAGGCACTGGCTAGAGAGCAATGTGAGCAATGAATCAAGTCTGCT	3180
Qy	3181	GTGCTCACCAAGTCCCATGTGACCCAGGCTGTGTGCTCAGAGGAGAGGGGTGCTTTTC	3240
Dp	3181	GTGCTCACCAAGTCCCATGTGACCCAGGCTGTGTGCTCAGAGGAGAGGGGTGCTTTTC	3240
Qy	3241	CTACCTTGCCAAAGGTGCTGTGTGGTTGGGGAAGTCCGACGTGCGGCTTTGTTTTCCCT	3300
Dp	3241	CTACCTTGCCAAAGGTGCTGTGTGGTTGGGGAAGTCCGACGTGCGGCTTTGTTTTCCCT	3300
Qy	3301	CTGCGCTCTTCTCTCTCTCTCAATATGTCTCATTTCTATCTCAACCAAGTTCCTAATGT	3360
Dp	3301	CTGCGCTCTTCTCTCTCTCTCAATATGTCTCATTTCTATCTCAACCAAGTTCCTAATGT	3360
Qy	3361	TCCTTGGGGATCCCACTGAGCTTTCATATACCTTCCCTCAGAGATCTCAACATCAACC	3420
Dp	3361	TCCTTGGGGATCCCACTGAGCTTTCATATACCTTCCCTCAGAGATCTCAACATCAACC	3420
Qy	3421	TTGGCTCGAAGAAATATCTATGCTGTGACACTGGATCTAGATCTAATCTTCTGAGCTCCA	3480
Dp	3421	TTGGCTCGAAGAAATATCTATGCTGTGACACTGGATCTAATCTTCTGAGCTCCA	3480
Qy	3481	GACATCTCTTTCATTTGATGTCTTACAGGCACTTAATTCAGCATCCTCCCAACTAA	3540

Db	3481	GACATCTTTTCCAAATGATGTTCTACAGGCACTTAAATTACGATCCCCAAACTTA	3540
Qy	3541	GCTTGGCATCTTCTTTACAAACCAACCTTCTCTGTGTTTCTGTCTTCAGTAATGAC	3600
Db	3541	GCTTTGCACTCTTTTCAAAACCAACCTTCTCTGTGTTTCTGTTTGAGTAATGAC	3600
Qy	3601	CCCAAAATGTGCTGATTACTACAAACCAAGTCACACAGGCTCATGATCTGGCCTT	3660
Db	3601	CCCAAAATGTGCTGATTACTACAAACCAAGTCACACAGGCTCATGATCTGGCCTT	3660
Qy	3661	GATTATCTTTCAGGTTTATCTCTCCCTGGCCCATTTCACTGTGTGCGACCCATACAA	3720
Db	3661	GATTATCTTTCAGGTTTATCTCTCCCTGGCCCATTTCACTGTGTGCGACCCATACAA	3720
Qy	3721	TCCTACATGAGGTTGAGACACACTGCTCTCTCATGTTTGGGCTTCGACATGCTCCCT	3780
Db	3721	TCCTACATGAGGTTGAGACACACTGCTCTCTCATGTTTGGGCTTCGACATGCTCCCT	3780
Qy	3781	GCTGTGAACACCCCTTCTCTCACTTGTGACCTGTGAAAATTCCTGTGATTTTTCAGCT	3840
Db	3781	GCTGTGAACACCCCTTCTCTCACTTGTGACCTGTGAAAATTCCTGTGATTTTTCAGCT	3840
Qy	3841	TGGGCCCCAATGCTTCTCTTGTGTGAAAACCTTCCCAACTTCTTACGACACTTAAG	3900
Db	3841	TGGGCCCCAATGCTTCTCTTGTGTGAAAACCTTCCCAACTTCTTACGACACTTAAG	3900
Qy	3901	CACCTGTCTATATTCCTCAGTGACACTTTCACCTACCTTGTGATGTCATGGCTAGG	3960
Db	3901	CACCTGTCTATATTCCTCAGTGACACTTTCACCTACCTTGTGATGTCATGGCTAGG	3960
Qy	3961	ATTGACGAGTCCCTTCTGTCTTGTGACAGTGAACCTTCTGAACTGAAAGAACAAGCTT	4020
Db	3961	ATTGACGAGTCCCTTCTGTCTTGTGACAGTGAACCTTCTGAACTGAAAGAACAAGCTT	4020
Qy	4021	GTTATCTCAGTGCCTCTCACAAATGCTGACATATAGTATTTCACTGACTGTCTT	4080
Db	4021	GTTATCTCAGTGCCTCTCACAAATGCTGACATATAGTATTTCACTGACTGTCTT	4080
Qy	4081	GGAATGATGAATGAATGAATTAATTAATGAAGAATGAATGAAGAATTAACGTATGGGT	4140
Db	4081	GGAATGATGAATGAATGAATTAATTAATGAAGAATGAATGAAGAATTAACGTATGGGT	4140
Qy	4141	ATTGACGAGTGAACAGTGTGAGATATGTTGTCAACTGATATGTTTGCAGATTAATGT	4200
Db	4141	ATTGACGAGTGAACAGTGTGAGATATGTTGTCAACTGATATGTTTGCAGATTAATGT	4200
Qy	4201	GCCACAGAGTGTCTGGGTACAGACTAGAGGCACTGTGTTTATGATATAGTACCTGGA	4260
Db	4201	GCCACAGAGTGTCTGGGTACAGACTAGAGGCACTGTGTTTATGATATAGTACCTGGA	4260
Qy	4261	TTTTCACAAACTGAGAGTGTATATGTGCAAAAGACACACATGTTGTCCACAGATGG	4320
Db	4261	TTTTCACAAACTGAGAGTGTATATGTGCAAAAGACACACATGTTGTCCACAGATGG	4320
Qy	4321	ACTGAGATGTGTAGGGCCACAGAAAGATATCTATAGCACATAGATAAAAAATGTGT	4380
Db	4321	ACTGAGATGTGTAGGGCCACAGAAAGATATCTATAGCACATAGATAAAAAATGTGT	4380
Qy	4381	GTAATATGCAAGTGTGAGATCTGTGGGATGCAAGTCAAAAAGAGATACTTTGAATGC	4440
Db	4381	GTAATATGCAAGTGTGAGATCTGTGGGATGCAAGTCAAAAAGAGATACTTTGAATGC	4440
Qy	4441	AGGGGACAAAGTGTGGGTATACCTCTCTGAAAAGAGAGAGATATCCCAAAATTCG	4500
Db	4441	AGGGGACAAAGTGTGGGTATACCTCTCTGAAAAGAGAGAGATATCCCAAAATTCG	4500
Qy	4501	TCCAAGATGAATTTCTGGAATCCCATCCCACTGATGCAAGTGTGGAATGTCAAGTG	4560
Db	4501	TCCAAGATGAATTTCTGGAATCCCATCCCACTGATGCAAGTGTGGAATGTCAAGTG	4560
Qy	4561	GAGAAATGCAAGTGAAGGCGTCTCTTCTTCAAGTCAAGTCAAGGCTGGAATTCACAA	4620
Db	4561	GAGAAATGCAAGTGAAGGCGTCTCTTCTTCAAGTCAAGTCAAGGCTGGAATTCACAA	4620

[illegible]

QY	5701	CATCTCTTCCACCTTGGAGCTCACTCTGACCTCTGGGCTCTTCTGTGTTCTCTCAACC	5768
Db	5701	CATCTCTTCCACCTTGGAGCTCACTCTGACCTCTGGGCTCTTCTGTGTTCTCTCAACC	5766
QY	5761	ATTCTCTCTCCAGGCGCATGGACAACAAGAAATGGGACTGGCCATGGGAACAGTACCCACA	5820
Db	5761	ATTCTCTCTCTCCAGGCGCATGGACAACAAGAAATGGGACTGGCCATGGGAACAGTACCCACA	5820
QY	5821	TGGCCCTGAAGTATATGCGCTGTAGCCCAACATCTAGTCTGTCTGACGTGTCGACAA	5880
Db	5821	TGGCCCTGAAGTATATGCGCTGTAGCCCAACATCTAGTCTGTCTGACGTGTCGACAA	5880
QY	5881	CCATGTGGCCACCTATGECCTTCAGTGTGAGAGATGCCCCCAATCCCCCAATGTGCTCT	5940
Db	5881	CCATGTGGCCACCTATGECCTTCAGTGTGAGAGATGCCCCCAATCCCCCAATGTGCTCT	5940
QY	5941	CACATCTCTTTACTTGTAATCTGCCACATCCCTGAACAATTTCTCATCTGTCACTGTGT	6000
Db	5941	CACATCTCTTTACTTGTAATCTGCCACATCCCTGAACAATTTCTCATCTGTCACTGTGT	6000
QY	6001	CACCTAATTTGTTGCCCTCTGTCCCACTCTTGACATGCCCTTCTGACATCCCTCATCTCT	6060
Db	6001	CACCTAATTTGTTGCCCTCTGTCCCACTCTTGACATGCCCTTCTGACATCCCTCATCTCT	6060
QY	6061	GAGGCATATTTCTCAATCTTGTCTGTCAAGGAGCCGAGCCCTTAACTTCACTCATCTCTCT	6120
Db	6061	GAGGCATATTTCTCAATCTTGTCTGTCAAGGAGCCGAGCCCTTAACTTCACTCATCTCTCT	6120
QY	6121	ACCATCTAATCCCAATGGCTGTGCCCCCTGTGGACCTCTGTGGGCGCCCTATGACCTCTGT	6180
Db	6121	ACCATCTAATCCCAATGGCTGTGCCCCCTGTGGACCTCTGTGGGCGCCCTATGACCTCTGT	6180
QY	6181	GTTTCTCCTTGCTCAATGCCCTGTGAGCCCTCTG6GACTCTGCCCTGTCCCTGACACTCTA	6240
Db	6181	GTTTCTCCTTGCTCAATGCCCTGTGAGCCCTCTG6GACTCTGCCCTGTCCCTGACACTCTA	6240
QY	6241	TGTGTCTCTGTACTCCTCTTG6GCTCCCTTGTCTTGCAATCTTTTCTGA6GTCTCTG6CT	6300
Db	6241	TGTGTCTCTGTACTCCTCTTG6GCTCCCTTGTCTTGCAATCTTTTCTGA6GTCTCTG6CT	6300
QY	6301	CCCCCTAATTTATCTGAGAACTCACTCTTGTCAG6TCTCG6TCTCTATGTGTCAGAC	6360
Db	6301	CCCCCTAATTTATCTGAGAACTCACTCTTGTCAG6TCTCG6TCTCTATGTGTCAGAC	6360
QY	6361	CCCTGGGCGATAGCACTGCTGGGGATGATATGTTCTCATTTGCTGAGAACAGCTGAGAG	6420
Db	6361	CCCTGGGCGATAGCACTGCTGGGGATGATATGTTCTCATTTGCTGAGAACAGCTGAGAG	6420
QY	6421	TGTTGGGTACTTTAGACTTTTAAAGG6GTGACTTCACTAG6CTCTGAGGTTTCTCTCTG	6480
Db	6421	TGTTGGGTACTTTAGACTTTTAAAGG6GTGACTTCACTAG6CTCTGAGGTTTCTCTCTG	6480
QY	6481	AGTATGCCAATGAGATACCCCTGCCCTTGAACCCG6GACTAATTTGGTAAAGCATCTTAA	6540
Db	6481	AGTATGCCAATGAGATACCCCTGCCCTTGAACCCG6GACTAATTTGGTAAAGCATCTTAA	6540
QY	6541	TAAATCCTAAGGGCTGTTCTGAGTTCA6GTCAAGGAGTAAATAGTCATCTGACAGTGTAG	6600
Db	6541	TAAATCCTAAGGGCTGTTCTGAGTTCA6GTCAAGGAGTAAATAGTCATCTGACAGTGTAG	6600
QY	6601	AATATCCCAAGAGAGTGAACAACAATCAATCCAACTGAGATATATGTATTAATTA	6660
Db	6601	AATATCCCAAGAGAGTGAACAACAATCAATCCAACTGAGATATATGTATTAATTA	6660
QY	6661	GGAACGTGTGAAGATATAAATCTGTAATATTTTTCACAACAATTTTGTGGCT	6720
Db	6661	GGAACGTGTGAAGATATAAATCTGTAATATTTTTCACAACAATTTTGTGGCT	6720
QY	6721	CTGACCCCTTGAGCAATTTTGAACAAGTATGACATATCAAGTTCTGTGAAAAAATACATCA	6780
Db	6721	CTGACCCCTTGAGCAATTTTGAACAAGTATGACATATCAAGTTCTGTGAAAAAATACATCA	6780

QY 6781 CCACATGAGAGCAAACTCCACAGAGATTGCACTATTAATAGAAACATACAGCTAA 6840
| | | | |
Db 6781 CCACATGAGAGCAAACTCCACAGAGATTGCACTATTAATAGAAACATACAGCTAA 6840
QY 6841 GATGAAACACACCTGTAGTAAATACAACTTAACTGAGAACATAGCCATAGTAA 6900
| | | | |
Db 6841 GATGAAACACACCTGTAGTAAATACAACTTAACTGAGAACATAGCCATAGTAA 6900
QY 6901 GAAACATAGATGTAAGAGAACACACAGCCATGTGTGGAGCCCATTTGGAGAGACACAC 6960
| | | | |
Db 6901 GAAACATAGATGTAAGAGAACACACAGCCATGTGTGGAGCCCATTTGGAGAGACACAC 6960
QY 6961 GACAAAGTAAATGACAGAAAGAGAGAGAGTGAAGATTTGTGAAACAGGGCCAC 7020
| | | | |
Db 6961 GACAAAGTAAATGACAGAAAGAGAGAGAGTGAAGATTTGTGAAACAGGGCCAC 7020
QY 7021 AGGAAACACACAGAAATAGAGAGAACACAGCCATCTAGAGATCACAGAACTTCATGG 7080
| | | | |
Db 7021 AGGAAACACACAGAAATAGAGAGAACACAGCCATCTAGAGATCACAGAACTTCATGG 7080
QY 7081 CCATGTGGCATTAATGAGATGCTACTGAACTCTTAAATGAAATATGCTATATGTTCC 7140
| | | | |
Db 7081 CCATGTGGCATTAATGAGATGCTACTGAACTCTTAAATGAAATATGCTATATGTTCC 7140
QY 7141 ATAGCTGTGAGAGAGCCCAACAGCATGAGAGAACCTTATATTAATAATCCAGGCC 7200
| | | | |
Db 7141 ATAGCTGTGAGAGAGCCCAACAGCATGAGAGAACCTTATATTAATAATCCAGGCC 7200
QY 7201 GGGCGTGTGAGTCAACGCTGTATCTTACCACTTTGGAGGCTGAGGCAAGTGTG 7260
| | | | |
Db 7201 GGGCGTGTGAGTCAACGCTGTATCTTACCACTTTGGAGGCTGAGGCAAGTGTG 7260
QY 7261 TTGAGGCGCTTGAAGCTTGAAGATTTGAGACACAGCTGGGCAACATGGCAAACTCATCT 7320
| | | | |
Db 7261 TTGAGGCGCTTGAAGCTTGAAGATTTGAGACACAGCTGGGCAACATGGCAAACTCATCT 7320
QY 7321 CTACAAATAATATTAATAATTAATGCGGGTGTGATGAGTGCCTTATAGTCCATCTACTT 7380
| | | | |
Db 7321 CTACAAATAATATTAATAATTAATGCGGGTGTGATGAGTGCCTTATAGTCCATCTACTT 7380
QY 7381 CAGAGGCTGAGCCCGGAGAGGCTTCAGTGAAGCCGTGATCGTGTACTGCACTCCA 7440
| | | | |
Db 7381 CAGAGGCTGAGCCCGGAGAGGCTTCAGTGAAGCCGTGATCGTGTACTGCACTCCA 7440
QY 7441 GCGTGGGTGACAGAGTGAAGCATGTCTCAAAAAAACAACAAACAAACAAACAAACAA 7500
| | | | |
Db 7441 GCGTGGGTGACAGAGTGAAGCATGTCTCAAAAAAACAACAAACAAACAAACAAACAA 7500
QY 7501 CAAACAAACAAACAAACAAACCAATATATATATATATATATATATATATATATATAT 7560
| | | | |
Db 7501 CAAACAAACAAACAAACAAACCAATATATATATATATATATATATATATATATATAT 7560
QY 7561 CTATTTTGTGTAATGACCAACATGACCCAGCTACAGATGGGGAGTCCCTCCCTCTC 7620
| | | | |
Db 7561 CTATTTTGTGTAATGACCAACATGACCCAGCTACAGATGGGGAGTCCCTCCCTCTC 7620
QY 7621 ACTGTGTAATTTTCTTCTCTGACTCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGAG 7680
| | | | |
Db 7621 ACTGTGTAATTTTCTTCTCTGACTCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGAG 7680
QY 7681 ATGAGATCTACCTGTGACCCAGGCTGAGTGCATATGGGCAATCTGTGTTACATGCA 7740
| | | | |
Db 7681 ATGAGATCTACCTGTGACCCAGGCTGAGTGCATATGGGCAATCTGTGTTACATGCA 7740
QY 7741 CCTGTGCTCTGTGGTTCAGAGCATCTCTGCTGAGCTCCCGTATAGTGGAGTAC 7800
| | | | |
Db 7741 CCTGTGCTCTGTGGTTCAGAGCATCTCTGCTGAGCTCCCGTATAGTGGAGTAC 7800
QY 7801 AGGCGCATACACCATGCTGCTGCTAATTTTGTATTTTGTGGGTACATGTACTAT 7860
| | | | |
Db 7801 AGGCGCATACACCATGCTGCTGCTAATTTTGTATTTTGTGGGTACATGTACTAT 7860
QY 7861 TTATTAATTTAATTTTGTATTTTGTATAGATAGAGGTTTCAACATGTTGGCAGGCTG 7920
| | | | |

Db 7861 TTATTAATTTAATTTTGTATTTTGTATAGATAGAGGTTTACACATGTTGGCAGGCTG 7920
| | | | |
QY 7921 GTCTGCAATCTGACCTGACGTGATCCGCTGCTGCGCTCCCAAGTGTAGATTA 7980
| | | | |
Db 7921 GTCTGCAATCTGACCTGACGTGATCCGCTGCTGCGCTCCCAAGTGTAGATTA 7980
QY 7981 CAGGCAATGAGCAACCGCTGGCCCTGATGAGTTTATATATCTCTTGGCTTTC 8040
| | | | |
Db 7981 CAGGCAATGAGCAACCGCTGGCCCTGATGAGTTTATATATCTCTTGGCTTTC 8040
QY 8041 ACAATTTGGCTTGCAGTGGACCATCATGTTCTCTCACTTCTGACTTCAATGATC 8100
| | | | |
Db 8041 ACAATTTGGCTTGCAGTGGACCATCATGTTCTCTCACTTCTGACTTCAATGATC 8100
QY 8101 TTTCACTCTCACTCTCACTGATCCTCCCTGAGTGTCTTTTCTTATAGATTTTC 8160
| | | | |
Db 8101 TTTCACTCTCACTCTCACTGATCCTCCCTGAGTGTCTTTTCTTATAGATTTTC 8160
QY 8161 CAGAGGGAATCTGAATGGCCCAATCCATTTTTCAGACCAACACATTAAGGTTT 8220
| | | | |
Db 8161 CAGAGGGAATCTGAATGGCCCAATCCATTTTTCAGACCAACACATTAAGGTTT 8220
QY 8221 GATTGCCAGCTATGATGATGCTACATTAATGGTTGGAACTCATTTACTCATTTG 8280
| | | | |
Db 8221 GATTGCCAGCTATGATGATGCTACATTAATGGTTGGAACTCATTTACTCATTTG 8280
QY 8281 CACAAAGCAGCATAGCTGTGTTCTCAAAATAGGCCCCCTGGGCAAGTGTGGCTCA 8340
| | | | |
Db 8281 CACAAAGCAGCATAGCTGTGTTCTCAAAATAGGCCCCCTGGGCAAGTGTGGCTCA 8340
QY 8341 TGCCTTAATCCCAACACTGTGGAGCCGAGGGGGGAGATTCATTTAGTCCAGAGTT 8400
| | | | |
Db 8341 TGCCTTAATCCCAACACTGTGGAGCCGAGGGGGGAGATTCATTTAGTCCAGAGTT 8400
QY 8401 CTAGACCAAGCTGGGCAACATGGTGAATCTCATCTCTAATAAATAACAAATAATAGC 8460
| | | | |
Db 8401 CTAGACCAAGCTGGGCAACATGGTGAATCTCATCTCTAATAAATAACAAATAATAGC 8460
QY 8461 CAGGTGTGTGACATGACCAAGTATGCTCCAGCTGTTCAAGAGGCTGAGGAGGATTTG 8520
| | | | |
Db 8461 CAGGTGTGTGACATGACCAAGTATGCTCCAGCTGTTCAAGAGGCTGAGGAGGATTTG 8520
QY 8521 CTCGAGTGTGGAGGAGATTTGCAATGAAACCTGACCTGTGCTTGCATTCACAGCTG 8580
| | | | |
Db 8521 CTCGAGTGTGGAGGAGATTTGCAATGAAACCTGACCTGTGCTTGCATTCACAGCTG 8580
QY 8581 GGTGACAGATTGAGACCTGTCTCAAAATAAATAATAATAATAATAATAATATGTTTC 8640
| | | | |
Db 8581 GGTGACAGATTGAGACCTGTCTCAAAATAAATAATAATAATAATAATAATATGTTTC 8640
QY 8641 CTGAGCAGGTAATTTCACTGAGGAAACCTCCAGGGGAGTGTGATGTCAGTCAACGCT 8700
| | | | |
Db 8641 CTGAGCAGGTAATTTCACTGAGGAAACCTCCAGGGGAGTGTGATGTCAGTCAACGCT 8700
QY 8701 GTATACTCAGTACCGGCTTAATAGAACTTGTGTATCAGCAGMAAACATAGGTATTT 8760
| | | | |
Db 8701 GTATACTCAGTACCGGCTTAATAGAACTTGTGTATCAGCAGMAAACATAGGTATTT 8760
QY 8761 ACTCAACAAATTTTGTGAGCATGATAGAGTGGGCAATTTGTCTAGGCACTGAGAT 8820
| | | | |
Db 8761 ACTCAACAAATTTTGTGAGCATGATAGAGTGGGCAATTTGTCTAGGCACTGAGAT 8820
QY 8821 ACAGTAGTCAACATGCAACAAGATGCTGCTGACAGGCTCTGCTAAAGTGAAGAG 8880
| | | | |
Db 8821 ACAGTAGTCAACATGCAACAAGATGCTGCTGACAGGCTCTGCTAAAGTGAAGAG 8880
QY 8881 GACATTAAGAAAGAAAGAAAGAAAGATTAATTTTGTATATTTAAAGGTTGTAAA 8940
| | | | |
Db 8881 GACATTAAGAAAGAAAGAAAGAAAGATTAATTTTGTATATTTAAAGGTTGTAAA 8940
QY 8941 GAAATTAAGACAGATAGTGGATAGAGGTGAGAGAAATGAGGGCTGTCTTGAAGAAA 9000
| | | | |


```

Db      8941 GAAATAAAGACAGATAGTGGGATAGAGTGAAGAGATGAGGCTGCTTCTGAGAAA 9000
Qy      9001 TGAATTTTGAAGTGAAGCTTCAATGATGAGAGAAATTAACCAAGATGCTGGAGGA 9060
Db      9001 TGAATTTTGAAGTGAAGCTTCAATGATGAGAGAAATTAACCAAGATGCTGGAGGA 9060
Qy      9061 AAAGCATTTTGGAGGGGTGAGCAGCATATCTTCAAGAAATCAAGAGAAAGCTGGTG 9120
Db      9061 AAAGCATTTTGGAGGGGTGAGCAGCATATCTTCAAGAAATCAAGAGAAAGCTGGTG 9120
Qy      9121 AGGCTGGAACACAGAGAAAGACAGCTGGGTGACTTGAAGGGGAGGAGCGCAGTGGCC 9180
Db      9121 AGGCTGGAACACAGAGAAAGACAGCTGGGTGACTTGAAGGGGAGGAGCGCAGTGGCC 9180
Qy      9181 AGGTTACCTAGACCTGGTAAAGGTTTCAACATTAAGAGAGATCAGAAAGTCTTGA 9240
Db      9181 AGGTTACCTAGACCTGGTAAAGGTTTCAACATTAAGAGAGATCAGAAAGTCTTGA 9240
Qy      9241 GAGGGCTGTGATATATTTCTAATCTATTTTATTAAGATCACTGACTTTTTCAGA 9300
Db      9241 GAGGGCTGTGATATATTTCTAATCTATTTTATTAAGATCACTGACTTTTTCAGA 9300
Qy      9301 ACATTAATTTAAAGTACAGACATGTTAGCAAGAAATCCAGTACGATCCGTGAGTT 9360
Db      9301 ACATTAATTTAAAGTACAGACATGTTAGCAAGAAATCCAGTACGATCCGTGAGTT 9360
Qy      9361 GTCCAAATTAGAGTGAATGACCGCTTGACATGAGATATGACAGAGAGGTGGTGA 9420
Db      9361 GTCCAAATTAGAGTGAATGACCGCTTGACATGAGATATGACAGAGAGGTGGTGA 9420
Qy      9421 TCACCATGATATATTTTGAAGGTAGAGCTGACAGCATTAATCTAATGCTAAGATAGCCG 9480
Db      9421 TCACCATGATATATTTTGAAGGTAGAGCTGACAGCATTAATCTAATGCTAAGATAGCCG 9480
Qy      9481 GGTGTGTGGCTTACCGCTGTATCTTGAACATTTTGGAGGCCAAGGCGAGTGCATACC 9540
Db      9481 GGTGTGTGGCTTACCGCTGTATCTTGAACATTTTGGAGGCCAAGGCGAGTGCATACC 9540
Qy      9541 TGAAGTCAGAGTTCGAGACAGCTTGACCAACATGCTGTAACCTGCTCTTAATAAAA 9600
Db      9541 TGAAGTCAGAGTTCGAGACAGCTTGACCAACATGCTGTAACCTGCTCTTAATAAAA 9600
Qy      9601 TACAAATTAAGCTGGGAATGCTGCAATGCTCTTAATCTCAAGCTTAATTTGGAGGCTGA 9660
Db      9601 TACAAATTAAGCTGGGAATGCTGCAATGCTCTTAATCTCAAGCTTAATTTGGAGGCTGA 9660
Qy      9661 GCGAGAGAGATCGCTTGAACCTGGAGGTGAATGTTGCACTGAGCGGAGATTGCACATT 9720
Db      9661 GCGAGAGAGATCGCTTGAACCTGGAGGTGAATGTTGCACTGAGCGGAGATTGCACATT 9720
Qy      9721 GCACTCAGGCTGGGGAACAGAGTGAATCTCGCTCTTAATTAATTAATGAATGAATG 9780
Db      9721 GCACTCAGGCTGGGGAACAGAGTGAATCTCGCTCTTAATTAATTAATGAATGAATG 9780
Qy      9781 ATATCAGTCAAGATGAGGAGAGGAGAAAGAGCTTCAAGATGACTGAGCTTTGCTGACT 9840
Db      9781 ATATCAGTCAAGATGAGGAGAGGAGAAAGAGCTTCAAGATGACTGAGCTTTGCTGACT 9840
Qy      9841 CAGCACTGAGTGGCTGGTGTGTTTGTCTTAATTTGGAGAAAGACTAGAGAGTGTGTG 9900
Db      9841 CAGCACTGAGTGGCTGGTGTGTTTGTCTTAATTTGGAGAAAGACTAGAGAGTGTGTG 9900
Qy      9901 TGTGTGTGGGGGAGAGAAATCAGTTTGGGCAATTAATGTTTGGGCTATTTGGACCC 9960
Db      9901 TGTGTGTGGGGGAGAGAAATCAGTTTGGGCAATTAATGTTTGGGCTATTTGGACCC 9960
Qy      9961 CATTAACATGTCAGTGAAGCAGCTGATTTGAGACCTTAAC 10000
Db      9961 CATTAACATGTCAGTGAAGCAGCTGATTTGAGACCTTAAC 10000

```

```

; Sequence 917, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 917
; LENGTH: 5234
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-917

```

```

Query Match      15.1%; Score 1509.4; DB 13; Length 5234;
Best Local Similarity 79.9%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 446; Indels 0; Gaps 0;

```

```

Qy      12 AGCAGAGCCAGCAAAATCTGTAGATTCAGACAGGTTCTGACACTGAGCAAGATTG 71
Db      3012 AGGTAGGTTTAAATTTTGAAGATTTAGATTAAGAGTTTGAATGAGTAAGTTG 3071
Qy      72 TTGAGGAATTTCTGATGAGATCATGGGCTCTAGAGAGGAATTAAGTTTCA 131
Db      3072 TTGAGGAATTTTGAAGATTTAGATTTAGGCTGTTTGAAGAGGAATTAAGTTTGA 3131
Qy      132 AGGCTGAGAGGGAAGAAAGAGAGAGGAGGAGCTTGAATGCTGCTCCATGCCCCA 191
Db      3132 AGGTTAGAGAGGAAGAAAGAGAGAGGAGGAGGAGTTTGAATGCTGTTTATTTGA 3191
Qy      192 CACCAGAAAGAGATGATCCCTGCAATGGGAGAGTGAATGAGACATTTGCTGTA 251
Db      3192 TATTTAGAAAGAGATATGTTTGTATATGGAGAGAGTGAATGATGATTTGTTGA 3251
Qy      252 GCGAGATGATGCTGCTGAGCTGCGAGAGCTCAGAGATCCAGCTTGCCTGACC 311
Db      3252 GTAGCATGATATTTGTTTGAAGTTTGAAGATTTAGAGATTTAGTTTGTATTAAT 3111
Qy      312 TATGAGAGAGGATATGTTTCAAGCAATTTTCAATGCTGAAGTGAAGAGATGGA 371
Db      3312 TATGAGAGAGGATATGTTTATATGATATTTTATTTGATGATGAGAGATATGA 3371
Qy      372 GCTGATGCAAGAGCTGCTGATGATGTTTCAAGAGTTCCGGAATGTTGTTTCT 431
Db      3372 GTTTGATGATGATGATGTTTGTATATGTTTGTATGAGTTTGGAAATGTTGTTT 3431
Qy      432 GTTGAAGAAATCTTGCAGAGTGAAGAAAGAGATCTGAGACTTTTGTGAATATAT 491
Db      3432 GTTGAAGAAATTTCTGTAGATGAGAAAGAGATTTGATTTTGTGAAGATATATAT 3491
Qy      492 GGGATGTAAGGGGTCTGAGCAATCTGAGAGGATCAGAGGCTTCAAGCTTGGCTAG 551
Db      3492 GGGATGTAAGGGGTCTGAGCAATCTGAGAGGATCAGAGGCTTCAAGCTTGGCTAG 3551
Qy      552 GAGAGAGGCTCTGAGCAATCTGAGAGGATCAGAGGCTTCAAGCTTGGCTAG 611
Db      3552 GAGAGAGGCTTCTGAGCAATCTGAGAGGATCAGAGGCTTCAAGCTTGGCTAG 3611
Qy      612 TGAATCACTAGCAAAATCTCAGACAGAAATGGGAGCTTTTGGAGTGGCCAGAGAT 671
Db      3612 TGAATTAATGATTAATTTATATATGAATGGGAGTGGTTTGGAGTGGCTTGAAGAGT 3671

```


QY	672	ACTGAGGATACAGAGTGTGATCCACGAGAAAGATGGAATGTGGGGCCGAGACACTGGAGAG	731
Db	3672	ATTGAGGATAGTAGGTGAGATTTTATGAGAGAGATGGAATGTGGGGCCGAGATTTATGGAGAG	3731
QY	732	AGAAACAGAGACTGTCAATTAAGAGGGCGCTGTGACTCCTAGACTCATTAATGCTACTAC	791
Db	3732	AGAAATAGGATTTGTATGATTAAGAGGGCGCTGTGATTTTAAATTAATTAATTAATTAAT	3791
QY	792	CATAACCTACCCCAATTCCTTAATTTCTCCTACCTAGAGGGGGGAAATTTGCAGAA	851
Db	3792	TATAATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAGGGGGGAAATTTGTAGAA	3851
QY	852	TTTGGCTGCACACTAGCAACACTACTCAGTACTGAAATCATTTTGTGCAATTTTTTCA	911
Db	3852	TTTGGTGTATATTTAGTAAATTTATTAATTAATTAATTAATTTTGAATTTTGTATTTTTTTTA	3911
QY	912	TTCAACAATAATTTCTGGACAACCTCTTAATTCGACGGCACTATTATTAAGATCAGGAT	971
Db	3912	TTTATATTAATTTTGTGGAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAGGAT	3971
QY	972	ATTATATGCTTAAACAGACAGGCAAAACAAAGCAAGCAACAACCATCACAGATTA	1031
Db	3972	ATTATATGCTTAAATTAAGTATAGTAAATTAAGTAAATTAATTAATTAATTAATTAATTA	4031
QY	1032	GTAGACAGATGAAGAAATTTCAAGTTTATAGTAAATTAATTAATTAATTAATTAATTAATTA	1092
Db	4032	GTAGATAGATGAAGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAGGTTGA	4092
QY	1092	ATGCTATGATTAAGCGGCTCAAGAAAGCTTCATTGAGAAAGTATGATTTAAACAGAGTCT	1151
Db	4092	ATGCTATGATTAAGCGGCTTAAAGAAAGCTTAAATTAAGAAAGGATTAATTAATTAAGTATGAGACTT	4151
QY	1152	AGCTAGAAATATTTGGAATTCAGATTCAAGTTCAATTTGTTCTGGTGGTTAAATTA	1211
Db	4152	AGTTAGAAATATTTGGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTA	4211
QY	1212	GCTTTTTCCTCCCAAGGTGAAACTACAAAGAAAGCTAATTAATTAATTAATTAATTAATTAATTA	1271
Db	4212	GTTTTTTTTTTTAAAGTGTGAATTAATTAAGAAAGCTAATTAATTAATTAATTAATTAATTAATTA	4271
QY	1272	TCTGGAAGAGAGACCCCTGTTCTGCTCACTACTGTCACCCCTCACTCCAGGCA	1331
Db	4272	TTTGGAGAGAGATTTTGTATAGTTAGGAAAGGCGAGAGGTTGGGTATGGGGTTGGATTA	4331
QY	1332	CTTTTTCGAAAGCCCTTTCAGAGTCAGAGAAAGCGAGAGCGTGGGCATGGGGCTTGGACA	1391
Db	4332	TTTTTTTGTAAAGTTTGTATAGTTAGGAAAGGCGAGAGGTTGGGTATGGGGTTGGATTA	4391
QY	1392	TTTGAACAACAGTGAACATTAATTTGCCCAAGCTACTAAGCCCAAGGTTAAAGCTGAAGA	1451
Db	4392	TTTGAATAATTAAGTGAATTAATTTTGTATTTTAAAGTTAATTAAGTTTAAAGGTTAAAGTTGAAGA	4451
QY	1452	GGCTTTGGGCAATGCCCCAGAAAGGCCCTGATGAAGCTTGAAAAAAGCTGTTCTCGAATTA	1511
Db	4452	GCTTTGGGTAATTTTGAAGAAAGCTTTTGAATGAAGTTTGAAGAAAGCTGTTTGTGAGTA	4511
QY	1512	TTTCTAAGTAAGTTATCTGTGTGTGTGTTACTTAAAGTAGTAAGTATTCGTCTCTTA	1571
Db	4512	TTTTTAAAGTAAGTTATTTGTGTGTGTGTTTAAAGTAGTAAGTATTTGTGTGTTTAA	4571
QY	1572	GCTGCTTAGAGCAAGGCGCTGACACAGTACACAGCAATTAAGTTCCTCCTCTTCTAC	1631
Db	4572	GTTGTTTAAAGTAGAGGCTTGAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA	4631
QY	1632	CTCCCCCATTTGAGATTAATCTCAATCACAAAAGTATCCTCAGTCTACTCACTTCCC	1691
Db	4632	TTTTTTTATTTGTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT	4691
QY	1692	TGACTTAATGATGCTGAGACCATTTGCCAGTGTGAGATCAACGCTGAGCGTCAAGACTG	1751
Db	4692	TGATTTAATGAGATTTTGAATTTATTTGTATGAGTGTGAGATTAATTAATTTGAGCGTTAGTATGTG	4751

Oy	1752	TAGCCCACTAGCTACGCTCTGAAAATTTGCTGAAAGGGGGGTGTGGGGGACAGCTGCACGGGAAAAG	1811
Dd	4752	TAGTTTAGATTATGTCTTGAAAATTTGTGAAAGGGGGTGTGGGGGAGTGTCTCGGAAAAG	4811
Oy	1812	GAGCTTGAGATTGAGATTTCTGTCCAGACCCTGACCTTATTTGCAGTGATGTAAATCAGCC	1871
Dd	4812	GAGTTTTGGAATTTAGATTTTGTATTGATTTGATTTTGAATTTTATTTGATGATGATTAATAGTT	4871
Oy	1872	AATATTGGCTTAGTCTCTGGAGACAGACATTTCCAGTAGACTTGGAGTGTGGGGTGTGTG	1931
Dd	4872	AATATTGTTAGTGTGGAGATGATATATTTTACTAGAGTTGGAGGTGGGGGTGTGTG	4931
Oy	1932	CTGTGTGCAACTATATTAGGGAGTTCACTGTGTCACCAGAGCTGTACGTGGCCCTGTG	1991
Dd	4932	TGTGTTTAATTTTAAATAGGAGATTTAATTTGTTATTGATGAGTGTGTTGTGTGTTTGTG	4991
Oy	1992	CAGCTCAGCANTGGCTAGGAGTACTGGAGACACCGTTGCATCTGGGGTGTGGAGCCTATGC	2051
Dd	4992	TAGTTTAGTATGTGTAGGATATTGGAGATATTCGTTGATTTGGGGTGTGTGAGTTTATGT	5051
Oy	2052	TGCTCTGTGGCCATTGGCAACCCCTTTCCTCCGTAGTAAAGCTGGAGCTAGAACGCAAG	2111
Dd	5052	TGCTTTTGTGTAATGTATATTTTCTTTTTCGTGAGTAAAGTTGGGATTAGAGCGAAG	5111
Oy	2112	GATTGAGTCTGGGCTTAGGGTAGGTAGGTAGGGCAGATTTTAGCGCTCGGTCAAAATTTGGGG	2171
Dd	5112	GATTGAGATTTTGGGCTTAGGGTAGGTAGGGTGTAGTTTATAGTTTCCGTTTAAATTTGGGG	5171
Oy	2172	TCAGGGGCTATGGGAAAGGAGTGGGTCCCAATGGATCAAGATATCTAATTTGTCTCCCT	2231
Dd	5172	TTTAGGGGTATAGGAAAGGAGTCGCTTTTATATGATTAAGATTAATTTGTGTTTTTTT	5231
Oy	2232	AGG 2234	
Dd	5232	AGG 5234	

```

US-10-311-455-918/c
RESULT 3
; Sequence 918, Application US/10311455
; Publication No. US2003014366A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PISENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis Associated with the Immune System by Determining
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 918
; LENGTH: 5234
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-918

```

Query Match	12.8%	Score 1281.6	DB 13	Length 5234
Best Local Similarity	73.4%	Pred. No. 0		
Matches 1638	Conservative 0	Mismatches 554	Indels 0	Gaps 0
QY	1	TCCTCTTCCCGAGCGCCGCAAAATCTGTGAGATTCAAGACGGGTTCTGACAGCTG	60	
DB	2234	TCCTCTTCCCGCAAAACCCGCAAAATCTATATAATTCAGAACAAATTTCAACCACTA	2175	

QY 61 AAGCAAGTTGTGAGAAATTCCTGATGAGATCATGGGTCCTCAGAGGAGATA 120
DB 2174 AAAACAAATATATATAAAATTCCTAATATAAAATCTATAAATCAAAAAATA 2115
QY 121 TAAGTTTCAGAGGCTGAGAGGAAAGAAAGTGAGGGGAGCTTAGAATGTGCTC 180
DB 2114 TAAAAATTCAGAAATATATAAAATATAAAATATAAAATCTTAAATATATATC 2055
QY 181 CCATTCGCCAACCAGAAAGAGACATGCTCCGCAATGGGGAGAGGTGATGATAGA 240
DB 2054 CCATTCGCCAACCAGAAAGAGACATGCTCCGCAATGGGGAGAGGTGATGATAGA 1995
QY 241 CATTCGCTGAGAGGAGATGCTTCGCCAGCTGCCAGAGCTCAGAGAGTCCAGCTT 300
DB 1994 CATTAATATATACAAACATATACCTTACCAATCCAAAACTCAAAAAATCCAACTT 1935
QY 301 GCCCATGACCTATGAGAGGAAATGATGTCACAGCAATTTTCATCGTAAGTCAGA 360
DB 1934 ACCCATTAACCTATATAAAATATAATATATTCACACATTTTCATTCGTAATCAAA 1875
QY 361 GAGGACATTCAGCTGATGAGAGGCTGCTGATCATGTTGTCAGAGGTCCGGAATG 420
DB 1874 AAAACATTAACCTATATAAAATCTAATATATATATTCCTCAAAAATTCGAAATA 1815
QY 421 TGTGTTTCTGTTGAGAGAACTTGCAGATGAAAGGATCTGAGACTTTGGTA 480
DB 1814 TATATTTCTTAATATAAAATCTTGCAGAAATATAAAATCTTAATACTTTAATA 1755
QY 481 AGATTAATATGAGAGCTGAGAGGCTGAGAGCTGAGAGCTGAGAGGAGATCAGGCTTCA 540
DB 1754 AATTAATATATAAAATCTAATAAAATCTAATAAAATCTAATAAAATCTAATAAA 1695
QY 541 GCCTTCGCTGAGAGGAGGCTGCTGAGACTTCACTGAGCTGAGAGTCTGAGCA 600
DB 1694 ACCCTAATATAAAATCTAATAAAATCTAATAAAATCTAATAAAATCTAATAAA 1635
QY 601 TAATTTCTTTCTGACTCATAGGCAATCTCAACAGAAATGGGGCACTTTGGAGTGG 660
DB 1634 TAATTTCTTTCTGACTCATAGGCAATCTCAACAGAAATGGGGCACTTTGGAGTGG 1575
QY 661 GCCCAGAGATCTGAGATGAGAGGATGCCAGAGAGATGAGATGGGGGCGGAG 720
DB 1574 ACCCAGAGATCTGAGATGAGAGGATGCCAGAGAGATGAGATGGGGGCGGAG 1515
QY 721 ACACTGAGAGAGAAACAGAGCTGTCAGATTAAGGGGCTGTCGACTCTGATCTGAT 780
DB 1514 ACACTGAGAGAGAAACAGAGCTGTCAGATTAAGGGGCTGTCGACTCTGATCTGAT 1455
QY 781 ATGCTTACTACATACCTACCTCCCAATCTCTAATATTTCTCTACCTGAGGGGGGAA 840
DB 1454 ATGCTTACTACATACCTACCTCCCAATCTCTAATATTTCTCTACCTGAGGGGGGAA 1395
QY 841 ATGCTGAGAAATTTGGTGAACCTGACCACTACTAGTACTGTAATGCAATTTTG 900
DB 1394 ATGCTGAGAAATTTGGTGAACCTGACCACTACTAGTACTGTAATGCAATTTTG 1335
QY 901 CATTTTTCATTCACAAATATTTCTGAGCAACTCTTATATGCGCAGCACTATTTAG 960
DB 1334 CATTTTTCATTCACAAATATTTCTGAGCAACTCTTATATGCGCAGCACTATTTAG 1275
QY 961 GAGTCAGGATATATATGTTAAACAGACAGGCAAAACAGAAAGCAACACCA 1020
DB 1274 AATTCAGAAATATATATATATATAACAAACAAACAAACAAACAAACCA 1215
QY 1021 TCACCAATATAGTACAGATGAAAGATTTCAAGTTTCTAGTAAATATAACAGAGC 1080
DB 1214 TCACCAATATATAGTACAGATGAAAGATTTCAAGTTTCTAGTAAATATAACAGAGC 1155
QY 1081 AAGGCTCTGAATGCTAGATTAAGGCTGAGAGAGGCTTCAATGAGAGTGAATTT 1140
DB 1154 AAAAAATCTAATATATATATATATATATATATATATATATATATATATATAT 1095
QY 1141 AAGCAGAGTCACTGAGAAATATTTGTAATTCAGATTACAGTTCTATTTGTCGGGTT 1200

DB 1094 AAAACAAATCAACTAATAATATATATAATTCATTAATCAATTCATTAATTCATTAAT 1035
QY 1201 GGTAAATATAGCTTTTCCCAAGTGAGAACTACCAAGAAAGCTAATATCTAGTAG 1260
DB 1034 AATTAATATATATATATATATATATATATATATATATATATATATATATATAT 975
QY 1261 TGGTGTGCTCTGTGAGAGAGACATCTCTGTTTCTGCTCATTAAGTCAACCTTC 1320
DB 974 TAATTAATCTCTTAATAAAACACATCTCTATTTCTACTCATTAATCTCAACCTTC 915
QY 1321 ACTTCCAGGACCTTTTGGAAAGCCCTTGGCAGTCAAGGAAAGGAGGCTGGGATG 1380
DB 914 ACTTCCAGGACCTTTTGGAAAGCCCTTGGCAGTCAAGGAAAGGAGGCTGGGATG 855
QY 1381 GGGCTTGACATTTGACACAGTGAAGATTTATTTGCCAGACTGACCTGAGGAGGT 1440
DB 854 AATCTTAACATTTAAACACATTAATAATTAATTAATTAATTAATTAATTAATTA 795
QY 1441 AAGCTGAAGAGGCTTGGGATGCCCCAGAAAGGCTGATGATGAGTGAAGAAAGCTG 1500
DB 794 AAAACATTAATAAACTTAACATATCCCAAAAAAACCCTAATATAAATTTAAAACTA 725
QY 1501 TTCTGAGATTTCTTAAGTATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
DB 734 TTCTGAGATTTCTTAAGTATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 675
QY 1561 TGTGCTCTGAGCTGAGCTTGAAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1620
DB 674 TACTATCTTAATCTTAATCTTAATAAACTTAACACATTAACACATTAATTAATTCCT 615
QY 1621 CCTTTTCTACCTCCCAATGTTGAGATTAACCTCAATCAAAAGGTGATCTCACTCT 1680
DB 614 CCTTTTCTACCTCCCAATGTTGAGATTAACCTCAATCAAAAGGTGATCTCACTCT 555
QY 1681 ACTCACTTCCCTCACTTAAGATGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 1740
DB 554 ACTCACTTCCCTCACTTAAGATGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 495
QY 1741 CGTCAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 1800
DB 494 CGTCAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 435
QY 1801 CGGAGAAAGAGGCTTGGATTCAGATTTCTGTCAGAGCTGACCTTAATTTGACATGA 1860
DB 434 CGGAGAAAGAGGCTTGGATTCAGATTTCTGTCAGAGCTGACCTTAATTTGACATGA 375
QY 1861 TGTAAATCAAGCAATTTGGTGTAGTCTGAGAGACAGACATTTCCAGTGAAGTGAAG 1920
DB 374 TGTAAATCAAGCAATTTGGTGTAGTCTGAGAGACAGACATTTCCAGTGAAGTGAAG 315
QY 1921 TGGGGGTGCTGCTGAGCACTCTAATATGAGGATTCAGAGGCTGAGAGGCTGAGAGG 1980
DB 314 TAAAAATATATCTACTACCACTCTAATATTAATAAAATTTCAACTAATACCCAAACATATCC 255
QY 1981 TGTGCTCTGACGCTCAGCATGCTAGGATCTGAGAGACAGGCTGAGAGGCTGAGAGG 2040
DB 254 TATTAATCTGACGCTCAGCATGCTAGGATTCAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 195
QY 2041 GAGGCTGAGGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2100
DB 194 GAGGCTGAGGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 135
QY 2101 TAGAGGAAAGATGAGATTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2160
DB 134 TAAAAAGGAAAGATTAATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 75
QY 2161 CAATTTTGGGCTCAGAGGCTATGAGGAAAGGATCGGTCCCAATGATCAAGATATCTATT 2220
DB 74 CAATTTTAAATTAATAAATCTAATAAATAAATAAATCAATCCCAATTAATCAAAATATCTATT 15
QY 2221 TTGTTCTCCCTA 2232

FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(480)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31405

Query Match 4.4%; Score 441.6; DB 11; Length 480;
Best Local Similarity 99.1%; Pred. No. 9.5e-104;
Matches 444; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4908 TTGACCTTATATATCTTTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 4967
DB 32 TGTGCTTTATATATCTTTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 91
QY 4968 GTACACTGAGTTTGTCTCAGTGTGGAATCTTCTCTCTGAGGTTCCATCTCTCT 5027
DB 92 GTACACTGAGTTTGTCTCAGTGTGGAATCTTCTCTCTGAGGTTCCATCTCTCT 151
QY 5028 TGTGTGCTTTATATCTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 5087
DB 152 TGTGTGCTTTATATCTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 211
QY 5088 TTCCAGACTCTCCCACTGCTTGTGAGCTGATCTGTGTCTTGTCTGCTGAAATT 5147
DB 212 TTCCAGACTCTCCCACTGCTTGTGAGCTGATCTGTGTCTTGTCTGCTGAAATT 271
QY 5148 TGTGTATATGACACCTCTCTTGAACCTCTCTGATGCTGAGCTCATGTGTCTCAG 5207
DB 272 TGTGTATATGACACCTCTCTTGAACCTCTCTGATGCTGAGCTCATGTGTCTCAG 331
QY 5208 TACCATTTATCTGAGCCATCTCTGAGACCCAGAGAAAGCAAAAGAGGCGTAACCCGCTC 5267
DB 332 TACCATTTATCTGAGCCATCTCTGAGACCCAGAGAAAGCAAAAGAGGCGTAACCCGCTC 391
QY 5268 TCACCAAAATGCTGTGATGATTGACCAAGGTGACCGGAGTGTCTGGAATCTTGC 5327
DB 392 TCACCAAAATGCTGTGATGATTGACCAAGGTGACCGGAGTGTCTGGAATCTTGC 451
QY 5328 TACGGGAACCATGAAGAGGCTTCTG 5355
DB 452 TACGGGAACCATGAAGAGGCTTCTG 479

RESULT 7

US-09-918-995-31407
Sequence 31407, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31407
LENGTH: 494
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(494)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31407

Query Match 4.4%; Score 441.4; DB 11; Length 494;
Best Local Similarity 99.8%; Pred. No. 1.1e-103;
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4908 TTGACCTTATATATCTTTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 4967

DB 52 TGTGCTTTATATATCTTTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 111
QY 4968 GTACACTGAGTTTGTCTCAGTGTGGAATCTTCTCTCTGAGGTTCCATCTCTCT 5027
DB 112 GTACACTGAGTTTGTCTCAGTGTGGAATCTTCTCTCTGAGGTTCCATCTCTCT 171
QY 5028 TGTGTGCTTTATATCTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 5087
DB 172 TGTGTGCTTTATATCTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 231
QY 5088 TTCCAGACTCTCCCACTGCTTGTGAGCTGATCTGTGTCTTGTCTGCTGAAATT 5147
DB 232 TTCCAGACTCTCCCACTGCTTGTGAGCTGATCTGTGTCTTGTCTGCTGAAATT 291
QY 5148 TGTGTATATGACACCTCTCTTGAACCTCTCTGATGCTGAGCTCATGTGTCTCAG 5207
DB 292 TGTGTATATGACACCTCTCTTGAACCTCTCTGATGCTGAGCTCATGTGTCTCAG 351
QY 5208 TACCATTTATCTGAGCCATCTCTGAGACCCAGAGAAAGCAAAAGAGGCGTAACCCGCTC 5267
DB 352 TACCATTTATCTGAGCCATCTCTGAGACCCAGAGAAAGCAAAAGAGGCGTAACCCGCTC 411
QY 5268 TCACCAAAATGCTGTGATGATTGACCAAGGTGACCGGAGTGTCTGGAATCTTGC 5327
DB 412 TCACCAAAATGCTGTGATGATTGACCAAGGTGACCGGAGTGTCTGGAATCTTGC 471
QY 5328 TACGGGAACCATGAAGAGGCTT 5350
DB 472 TACGGGAACCATGAAGAGGCTT 494

RESULT 8

US-09-918-995-31269
Sequence 31269, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31269
LENGTH: 478
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(478)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31269

Query Match 4.4%; Score 440.8; DB 11; Length 478;
Best Local Similarity 99.5%; Pred. No. 1.5e-103;
Matches 442; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4908 TTGACCTTATATATCTTTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 4967
DB 34 TGTGCTTTATATATCTTTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 93
QY 4968 GTACACTGAGTTTGTCTCAGTGTGGAATCTTCTCTCTGAGGTTCCATCTCTCT 5027
DB 94 GTACACTGAGTTTGTCTCAGTGTGGAATCTTCTCTCTGAGGTTCCATCTCTCT 153
QY 5028 TGTGTGCTTTATATCTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 5087
DB 154 TGTGTGCTTTATATCTTACCTTAACCTAAGCTCCAGAAACCTATGCTATCTCT 213

Db	11815	-----ACAGTGGCTCAGCCCGCTAATCCAGCAATTTTGGAGGTTCAAGCAGGTTGATAA	11865
QY	8385	CTTGAGTCCAGGAGTTTGAACACAGCCTTGGCAACATGGTGAATCTCATCTTACTAA	8444
Db	11870	TTTATAGGTCGGAATTTGAGACCAAGCCTTGCACACATGGGGAAAACCTGTCTTACTTAA	11929
QY	8445	AATACAAAAAAATTAGCCAGGTGTGGTGGCAATGCACCAATGATCCCACTGTTCCAGGAGCC	8504
Db	11930	AATACAAAAATTAGCTAGGCAGATGTGGTGCACACCTGTAACTCCAGCTACTTGGGGGAC	11989
QY	8505	TGAGTGGGAGGATTTGCTCGAGTGTGGGAGGACGAAATTGCATGAAACCTGATCTTGCC	8564
Db	11990	TGAAGCAGGAAATCTTTGAACCTGGAGAGCGAGGTTGTGTGAGCCGAGATCAATCC	12049
QY	8565	TCTGCAATCCAGCCTGGGTGACAGATTTGAGACCCCTGTCTCAAAAAACAATAAATTAAT	8624
Db	12050	ACTGCATCCAGTCTGGGTGACAGAGTGAACCCCTATCCAAAAAATTAAAAAATAAAAA	12109
QY	8625	AAA 8627	
Db	12110	AAA 12112	

```

RESULT 10
US-10-017-161-1795/c
: Sequence 1795, Application US/100017161
: Publication No. US2003014368A1
: GENERAL INFORMATION:
: APPLICANT: SUMA, MAKIRO
: APPLICANT: ASAI, KIYOSHI
: APPLICANT: AKIYAMA, YUTAKA
: APPLICANT: ABRATANI, HIROYUKI
: TITLE OR INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
: FILE REFERENCE: 084335/0152
: CURRENT FILING DATE: 2002-12-18
: PRIOR APPLICATION NUMBER: JP 2001/246789
: PRIOR FILING DATE: 2001-06-18
: NUMBER OF SEQ ID NOS: 2430
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1795
: LENGTH: 43419
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURES:
: NAME/KEY: source
: LOCATION: (1)..(43419)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (201)..(278)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (8419)..(8609)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (110482)..(10594)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (119132)..(19408)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (41900)..(42106)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (42300)..(42393)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (43094)..(43219)
: FEATURE:
: NAME/KEY: modified base
: LOCATION: (20434)..(20533)
: OTHER INFORMATION: a, t, c, g, unknown or other
: US-10-017-161-1795

```

Query Match	3.7%;	Score 372;	DB 13;	Length 43419;
Best Local Similarity	56.6%;	Pred. No. 2.2e-84;		
Matches 852;	Conservative 0;	Mismatches 535;	Indels 119;	Gaps 10;

OY	7176	ACCTTATATTAAAAAATCCAGGCGGGCGTGGTAGTCACGCCTTAATCCTAGCACTT	7235
Dd	21360	ACCOCATCTAAATAATCAAACAATGGCTGGGCCGGTGGCTCACACTGCACTACTAGCACTT	21301
OY	7236	TGGAGGCTGAAGCAGGTGGATTGCTTAGCGGGCTTTGAGCCTTAGAGATTGAGACCAAGC	7295
Dd	21300	TGGAGAGCTGAAGCAGGTGGAT-----CGCTTGAATCTCAGGAGITCAAGAACAAGCC	21258
OY	7296	TGGGCAACAAGGCAAAAACCTCATCTCTACAAAATAATA-----AAAATTA	7341
Dd	21249	TGGCAACAAGTGGTAAACCCCATCTATGCAAAAAAAAAAAAAAAAAAAATAACAATAATTT	21190
OY	7342	GTCGGGTGTGTGTAGTGCTTCTATATAGTCCCATCTACTTCAGAGCTGAGCCCGAG--	7399
Dd	21189	ACC GGAGATGTGTGGCAAGCATTTGTGTGTCTGCTACTTTGGAGAGCTGAGGGGGAAGAT	21130
OY	7400	-----GTGAGGCTTCAGTAGACCGCTGATCGTGCTACTGCACTCCAGCC	7443
Dd	21129	CGCTTGAAGCCAGGAGAGTCAAGATTGCAAGTGAACAGTGAATTGACACACTGCAATTCAGCC	21070
OY	7444	TGGGTGACAGTGAAGACATATCTCAAAAAAAAAACA AAAAACAAAAACAACAACA	7503
Dd	21069	TGGCGCAAGAGCAAGACCCTGCTCTAAATCAATAAATCAAACHCTTCTTGCCAA	21010
OY	7504	ACAAACAACA AAAAAACCCATATATATATATATATATACCTAGCTGAGGTGAGATGCATTA	7563
Dd	21009	TGTGTATCTCTCTCTGTTTTGATGATAGGCGTCAGACGCATCTAGATACCAAGCTAG	20956
OY	7564	TTTTGGTAAATAATACCAACAATGACCAAGCTACAGCATGGGAGAGTCCCTCCCTCTACT	7623
Dd	20949	AACCTTGCACCTTGTGATCCCTCTCTCTCTCTGTATCTAGAAACACTGCATCTGCACCT	20890
OY	7624	GGTAAATTTTTCTTCTCTGACTCACAGTTTTGTGTGTGTGCTGTGTTTGAGATG	7683
Dd	20889	ACT-----TGAGGCTGCTCTGACTGTGTGTTTTTTTTTTTTTTTTTTTTTGTAGATG	20839
OY	7684	GAGTCTCACTCTGTCAACCCAGGCTGAGTGCATATGGCGCAATCTTGTTCACTGCAACT	7743
Dd	20838	GAGTTTTGCTCTTTCACCCAGCATGGAATGGAAGTGCACATCTCAGCTCACTGACGCTG	20779
OY	7744	CTGCTCTCTGGGTTCAAAGCATCTCTCTGCTCAAGCTCCCGTANAGCTGGAGCTACAG	7803
Dd	20778	CTGCCCCCGGGTTCAAAGCAATTCCTGCTCAAGCTCAAGATGAGCTGGGATTAACAG	20719
OY	7804	CGCATTAACACATGCGCGCTAATTTTTGTGATTTTTTGGGTTAACAATGACTAATTTA	7863
Dd	20718	CGTGTCTTACCAACTGG-----	20700
OY	7864	TYAATTTAATTTTTGTATTTTTAGTAGAGATAGGTTTACCA-TGTTGGCCAGGCTGT	7922
Dd	20699	-----TYAATTTTTGTATTTTTTAGTACAGACAGGGTTTCAACATGTTGGCCAGGCTGT	20645
OY	7923	CTGCAATCTCTGACCTCAGGTATCCGCTGCTCGGCTCCCAAGTGTCTAGATTAACA	7982
Dd	20644	CTTGAACTCTCTGACCTCAGGTATCCGCTGCTCGGCTCCCAAGTGTCTAGATTAACA	20585
OY	7983	GGCATGAGCAACACGCGTGGCCCTCATAGGTTTATCTAATCTCTTGTGCTTCTTAC	8042
Dd	20584	GGCATGAGCACTGGCGCCAGCTGACTTCTTTTCTTTTCTTTTCTTTTCTNNNNNNN	20525
OY	8043	AAC TTTGGCTTGACGTGACATCATGTTCTCTCCACTTCTCACTTCACTCATGATCTT	8102
Dd	20524	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	20465
OY	8103	TCAGTCTCAGTTCACA CTGATCTCCCTCA-----GTGCTCTTTTTCTTACTAAGA	8156
Dd	20464	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTATTTATTTAATTTTGTGAGACAGGT	20405

```

QY 8157 TTTCCAGAGAGGGAATCTGAATGCCCCAGTCCATATTTTCAGACCAACACATTTAAAGT 8216
DB 20404 CTCTCTGTGGCCAGGCTGGAATGCAATGCTGATCTTGCTCACTGCACTCTGCGCT 20345
QY 8217 GGTGATGCGCAGCCTATGATGCTGCTACATTAATGGGTTGGAACTCATCTTTACTTC 8276
DB 20344 CCGGGTTGAAGGATCTCTGCTCCAGCTCCCAATGAGTGGAGCTTAACAAGCATGT 20285
QY 8277 ATTGCACAAAGCAGCTAGCTGCTGGTTC-AAAATAGGCCCCCTGGGCAAGTGGTG 8335
DB 20284 GCCACTGACCCGGCCAAATTTTAAATTTTAAATTAACAAAGTTGAGCGCCGACGTTG 20225
QY 8336 GCTCATGCTATATATCCCAACACTGTGGAGCGGAGGCGGAGAT-----8382
DB 20224 GCTCAACCTGTATCTCCAGCACTTTGGAGGCCAAGTGGGAGATCAACCACTGACTT 20165
QY 8383 -CACTTGAAGTCCAGAGTTCTTGAAGCAAGCTGGGCAACATGTGAATCTCATCTTACT 8441
DB 20164 CCACTGGAAGTCCAGACTTTTGAAGCAAGCTGGGCAACATGTGAATCTCATCTTACT 20105
QY 8442 AAAAAATACAAAAAATTAAGCAAGTGTGTGGATGACCACTGATCTCCAGCTTTACGA 8501
DB 20104 AAAAAATAC-AAAAATAGCCAAATGTGTGGCGGCCCTGTAGCCCAAGTACTCAGGA 20046
QY 8502 GGCTGAGTGGAGAGATTGCTCGAGTGTGGAGGAGAGATGTCAGTGAACGCTGACTGT 8561
DB 20045 GGCTAAGTGGAGAGATCACTTGAACCCAGAGGCTGAGGTTCCAGTGAAGCAAGTTGT 19986
QY 8562 GGCTGCAATCCAGCCTGGGTGACAGATTGAGACCCCTGTCTCAAAAAACAAATTAATA 8621
DB 19985 GCCACTGCACTCCAGCCTGGGTGACAGAGCAAGCTGTCTCAAAATTAATTAATAA 19926
QY 8622 AATPAA 8627
DB 19925 TACAAA 19920

```

RESULT 11

US-10-017-161-1795

```

/ Sequence 1795, Application US/10017161
/ Publication No. US20030143668A1
/ GENERAL INFORMATION:
/ APPLICANT: SUMA, MAKIKO
/ APPLICANT: ASAI, KIYOSHI
/ APPLICANT: AKIYAMA, YUTAKA
/ APPLICANT: ABURATANI, HIROYUKI
/ TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
/ FILE REFERENCE: 084335/0152
/ CURRENT APPLICATION NUMBER: US/10/017,161
/ PRIOR FILING DATE: 2002-12-18
/ PRIOR APPLICATION NUMBER: JP 2001/246789
/ NUMBER OF SEQ ID NOS: 2430
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1795
/ LENGTH: 43419
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: source
/ LOCATION: (1)..(43419)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (201)..(278)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (8419)..(8609)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (10482)..(10594)
/ NAME/KEY: CDS
/ LOCATION: (19132)..(19408)

```

```

/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (41900)..(42106)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (42300)..(42393)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (43094)..(43219)
/ FEATURE:
/ NAME/KEY: modified base
/ LOCATION: (20434)..(20533)
/ OTHER INFORMATION: a, t, c, g, unknown or other
US-10-017-161-1795

```

Query Match 3.7%; Score 368.2; DB 13; Length 43419;
 Best Local Similarity 59.5%; Pred. No. 2,2e-83;
 Matches 866; Conservative 0; Mismatches 478; Indels 111; Gaps 10;

```

QY 7269 CTTGAGCCTAGAGATTGAGACAGCCTGGGCAACATGSCAAACCTCATCTTACAAA 7328
DB 12442 CTTGAAAGCCAGAGATTGAGACAGCCTGGGCAACATGTTGAAACCCCATCTTACAAA 12501
QY 7329 AATATATAAATTAGTGGGTTGTAGTGCCTTCTATAGTCCATCTACTTACAGAGCT 7388
DB 12502 AATACAAAATTAGCAGGAGATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 12561
QY 7389 GAG-----CCGGAAGTGAAGCTTCAATGAGCCCTGATCTGTCTA 7430
DB 12562 GAGTCAGAGAAATCCCTTGAACAGGAGAGAGAGTTCAATGAGATGATCCGACCA 12621
QY 7431 CTGCACTCCAGCCTGGGTGACAGATGAGACCATGCTCAAAA-AAAAAACAACAAA 7489
DB 12622 CTGCACTCCAGCCTGGGCAACAGATGATGATCTGTCTCAAAAACAACAAAACA 12681
QY 7490 AAAAAACAACAAAACAACAAAACAACAAAACAACAAAACAACAAAACAACAAAACA 7529
DB 12682 ACAACCTGAGCCTGGGTATGATTAAGAAAGAGTTTAATGCTCAAGTTCTGAG 12741
QY 7530 --TATATATATATCTAGTGAAGTGAATGCACTATTTTGTAAATCAACCAATGAC 7587
DB 12742 GCTGTACAAAGATGACACAGCATCTGCTGGGCTTCTGGGAGGCTTCAAGAGCTTTTA 12801
QY 7588 CCAAGTACAGCATGGGCGAGTCCCTCCCTCCTCAGTGAATTTTT-----CT 7636
DB 12802 CTTTGGCAAAAGGGTGAAGGACACACATGCTTGAAGCAGAGCAAAAAGAGAGAG 12861
QY 7637 TTCTGACTCAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 7696
DB 12862 TGTGGGAGGTGACATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTG 12921
QY 7697 TCAACCAAGCTGAGTGAATGAGGCAATCTTGTGTTCACTGCAACTCTGCTCTGGGT 7756
DB 12922 TCGCCCAAGCTGAGTGAATGAGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12981
QY 7757 TCAAGGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7816
DB 12982 TCAAGGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13041
QY 7817 GCTGCTAATTTTGTATTTTCTTTTGGTTACAAATGACTATTATTAATTAATTT 7876
DB 13042 GCCCGGCTAAT-----GTTT 13057
QY 7877 TGTATTTTATGATAGATAGAGTTTCAAGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 7936
DB 13058 TGCATTTTATGATAGAGAGGAGGTTTCAAGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 13117
QY 7937 CTCAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7996
DB 13118 CTC--GTATTCGCGCCACTGAGCTTCCAAAGTCTGATTTACAGGGGTAGACACCG 13175
QY 7997 CGCTGGCCCTCATAT--GGTTTATCTATTTCTTTGCTTCTTCAACATTTGGCTT 8053

```


Db	13176	CGCCACGACGATGCGACATCTTTTAAACAACTAAGACTGTGAAAGAACTCATATCTCGA	13233
Oy	8054	GCAGTGGACCAT-----CANTGTCCTCCACTTTCTCACTACTTCAATGATCTTTCAATC	8108
Db	13236	GAACAGCACCAATCCATGACAGAAATCCATCCATGACCCAAACACTCCACCATGCGCA	13295
Oy	8109	TCAGTTCCAACTGATAC-----CTCCCTCAGTGTGCTCTTTTCTCTAGTAAGATTTCCAGA	8164
Db	13296	CCTCCAAACCTGGGGAATTACAATTCCACATGAGATTTGGGCAAGGACAAAGTATTAACAGT	13355
Oy	8165	GAGGGAATCTGAATGCCCCAGTCCATATTTTTCAGACCAACCAATTTAAAGTGTGATT	8224
Db	13356	GATAGGTAACGTCCACATACAGGTATATAGGGCTACCTGGAGGAGGCAACGCTGAAGTGA	13415
Oy	8225	GCCAGCCTATGATTTGGCTACATTAATGGGTTGGGAACTCATATTTATCTTCAATCACAC	8284
Db	13416	AGGCTTGCTGTGAAGCTATGGGGGCTTTTATGAGGAAGGCTCTATGCAAGGTGAGGA	13475
Oy	8285	AAGCAGCATAGCTCTGCTTCTCAAAATAGGCCCCCTGGGCGCAGGTGTGTGCTCATGCC	8344
Db	13476	TGGATATGGGGCTTGATGGTTAGAAATGGGAGAGGAGGACAGTCAAGTGTCAACACC	13535
Oy	8345	TATATACTCCAACTATGTTGGAGGCGGAGGGGGGCAATTCATTGAATCCAGAGTTCTTGG	8404
Db	13536	TGATATCCCGACATTTTGGGAGGCGCGAGGTGGGTGATCCCTGAAGTCAGGAGTTCCAG	13595
Oy	8405	ACCAGCTGGGCAACATGTGGAATCTCATCTCTCACTAAATAATACAAAAATTTAGCCAG	8464
Db	13596	ACCAGCTGGGTAAATGATGTGAATCCCACTCTCATATAAAAAAT-AAAAATTTAGCAGG	13654
Oy	8465	TGTGTGGCATGACCACTGATGTCCCACTGTTTTCAGAGGTGAGGTGGAGGATTTGCTCG	8524
Db	13655	CATGTGTGGCGTAACCTGTATATCCCACTACTCAGGAGGCGGAGGCAAGAAATTTGCTTG	13714
Oy	8525	AGTGTGGAGGCAAGATTTGCAATGGAACCGATGACTGTGCAATCCAGCTGGGAG	8584
Db	13715	TACCTGGAGGTAGATTTGCAATGAGCCAAAGATTTGTCCATTGCACTCCAGCTTGGGCG	13774
Oy	8585	ACAGATTGAACCTGTCTCAAAAAACAAATTAATTAATAATTAATTAATGTGTTCTCTGA	8644
Db	13775	ACAGAGGAGACCTCCGCTCAAAAAAATTTAAAAAATTTAAAAAAGTGGAGAGAGG	13834
Oy	8645	GCAAGGATATTTTCAG	8659
Db	13835	GAACGTTGACTGCAG	13849

```

RESULT 12
US-10-374-979-11
: Sequence 11, Application US/10374979
: Publication No. US20030219793A1
: GENERAL INFORMATION:
: APPLICANT: John P. Carulli et al.
: TITLE OF INVENTION: THE HILL BONE MASS GENE OF 11q13.3.
: FILE REFERENCE: 032786-021
: CURRENT APPLICATION NUMBER: US/10/374,979
: CURRENT FILING DATE: 2003-03-04
: PRIOR APPLICATION NUMBER: US 09/544,398
: PRIOR FILING DATE: 2000-04-05
: PRIOR APPLICATION NUMBER: US 09/543,771
: PRIOR FILING DATE: 2000-04-05
: PRIOR APPLICATION NUMBER: US 09/229,319
: PRIOR FILING DATE: 1999-01-13
: PRIOR APPLICATION NUMBER: US 60/071,449
: PRIOR FILING DATE: 1998-01-13
: PRIOR APPLICATION NUMBER: US 60/105,511
: PRIOR FILING DATE: 1998-10-23
: NUMBER OF SEQ ID NOS: 109
: SEQ ID NO 11
: LENGTH: 66933
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-374-979-11

```

Query Match	3.5%;	Score 355;	DB 13;	Length 66933;
Best Local Similarity	59.6%;	Pred. No. 7.8e-80;		
Matches 871;	Conservative 0;	Mismatches 460;	Indels 130;	Gaps 11;

OY	7184	TTAAAAATACCAGCGCGGCGGCGGTGAGTACAGCGCCGTATATCTGACATTTTGGGAGGC	7243
Db	61260	TTTAAAGACCTTTGGGCGCCAGCGTGAATGGCTTTGGAGCGTATATCTCAGCATTTAAGGAGGC	61311
OY	7244	TGAGGCGAGTGTGATGTGTTAGCGGGCTTAGAGCTTAGAGTTTGAAGCAGCGCTGGGCAC	7303
Db	61320	TGAATATGAGGGAACTGATAGG-----CCGGAGTTTATGAGACAGCGCTGGGCCAAC	61377
OY	7304	ATGCAAAACCTCATCTCTCAAAAAAATATATAAAATATAGTCGGGTGTGTAGTGCCTTC	7263
Db	61371	ATGAGAAAGAACCCGCTCTGTAGTAAAAATTTAAAAATTAGCCGGGTATGTGATCCAGCT	61433
OY	7364	TATAG-----TCCCATCTACTGAGAGCTGAGCCCGGAGAGTGTGAGGCTTCAGTAGCC	7418
Db	61431	ACTTGGGAGTCTGAGAGCATGAGAAATGCTTGAACTATGGGAGGCGGAGGTGCAGTAGCC	61499
OY	7419	GTATTCGTGTCTACCTGCACTCCAGCTCGGTGTGACAGAGTGAAGACATGTCTCAAAAAAAC	7478
Db	61491	GAGATCCGCGCATTTGCACTCCAGCGCTGGGGGACAGCGAGCATCTGTCTCAAAAAAAA	61555
OY	7479	AAAAACAAAAACAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC	7538
Db	61551	AAAAAAAAAAAAAGAAAAAGAAATATCAATCTCTCTTTTATGCAATATATATATATAT	61611
OY	7539	ACCTAGCTGAGGTGAGAAATGCACTATTTTGTGTAATTCACCAACATGACCCAGCTACAGC	7598
Db	61611	ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT	61665
OY	7599	ATGGGAGAGTCCCTCCCTCTCACTGTGTAATTTTCTTCTGTGCTCAAGTTTGT	7658
Db	61664	---AGGCTTCCCTGCTCTGATCATATAAAAAACATATATTTTACACTCTCTCTTTT	61720
OY	7659	GTTCGTGTGTCTGTGTGTGTGAGATGAGTCTCACT-CTGTCAACCAAGCGTGAAGTCAAT	7717
Db	61721	TT-----TTTGAACAAGTTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	61766
OY	7718	GGCGCAATCTTGGTTCACTGCAACTCTGTCTCTCGGTGAAGCGATCCTCTGCTCA	7777
Db	61767	GGCGCAATCTCACTCACTGTAACCTCCGCGCTCCGGGTGAGATGATTTCTCGCTTAA	61828
OY	7778	GCCCTCCGTAATGTGGGACTAGGCGGCACTAACAGACCGCGGTAATTTTGTATTT	7837
Db	61827	CTTCCCGAGTGTGGGATTAATAGGATGACACACATGCGCTGGCTAA-----	61877
OY	7838	TTTTTGGGTAAACATGATCTATTTATTTAATTTTGTATTTTGTATGAGA-TAG	7896
Db	61876	-----TTTTGTACTTTTGTATGAGACGGG	61899
OY	7897	GGTTTACCAATGTTGGCAGGCTGTGTCTGAACTCTGACCTCAAGTATTCGGCTGCT	7956
Db	61900	GGTTTCTCAATGTGTGACAGGCTGTGTCTCGAATCCGCACTCAGGTATCCACCACT	61958
OY	7957	CGGCTCCCAAGTGTGAGATTAACAGGCACTAGGCAACACAGCGTGGCCCTCATAGGT	8016
Db	61960	CGGCTCCCAAGTGTGAGATTAACAGGCACTAGGCAACACAGCGTGGCCCTCATAGGT	62018
OY	8017	TTTATCATATCTCTTGTCTTCTTCAAACTTTTGGCTTGAAGTGAACCATATGTTCTCT	8076
Db	62020	TCTTAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	62055
OY	8077	CCACTTTCATCTTCAATGATCTTTCAGTCTCACTGATCCAACTGATATCTCCCTCAGTT	8136
Db	62059	TAAATCTTCTGGGATGTAAATTAATTTGATGAGCAAGTGGGATCCAGCTGTGTTCTTG	62118
OY	8137	GCCTCTTTTCCAGTAATTTCCAGAGAGGGAATCTGAATGGGCCATGCAATTTTC	8196
Db	62119	GCTGATGGCTTAATGGGTGGGTGAATTAATGTCGGGGTCTATCAGAGAGCAAACTCTATG	62178

QY	819,9	AGACACACACACATTAAAGTGGTGAATGCGACGCTATGATATGGCTAATTAAAGGCTT	825,6
Db	6217,9	AG-----AAATTGAACAGAGAAAGTTCTCGTACAGGCTTATTTACAGGGACTGAA	6223,0
QY	825,7	GGGAACTCATCATTTTACTTCATTGCAAAAGCAGCATAGCTGTGTTCTCAAAATAGGAC	831,6
Db	6223,1	TAGACAAATTGAACAGTAGATGTACAGAACTCT-----AGAAATGCAAG	6227,7
QY	831,7	CCCTGGGCCCAGGTGTGTGTGGCTCATGTGCTTATATCCCAACATGTGGAGGCCAGAGGGG	837,6
Db	6227,8	GAATAGGCCAGGACATGTTGGCTCACACTGTCAATCCAGACCTTGGAGACCAAGGCCG	6233,7
QY	837,7	GCAGATCACTTGAATGCAGAGGTTCTAGACCAAGCCTGGGCAACATGTAATAATCATCT	843,6
Db	6233,8	GTGGATCACTGAGAGTACAGAGTTCCGAGACCAAGCTGGCCAACTATGTBAACCCATCT	6239,7
QY	843,7	CTACTATAAAATAC-AAAAAATTAGCCAGGTGTGTGGCATGCAACCATGTGTCCACGCTGT	849,5
Db	6239,8	CTACTATAAAATACAAAAAATTAGCTGGGTGTGTGGCCCATGCTGTATATCCACGCTAC	6245,7
QY	849,6	TCAAGAGGCTGAGGTGGAGATGTGCTCGAGTGTGGAGGCGAGAGTTGACAGTGAACCGT	855,5
Db	6245,8	TCCGAGAGGCTGAGGCGAGAGAAATCACTTGAACCTGGAGAGCGAGGTTCAAGTAGCCGA	6251,7
QY	855,6	GACTGTGCTCTTGCATTCAGACCTGGGTGAACAGATTGAAGCCCTGTCTCAAAAAACAAT	861,5
Db	6251,8	GATCATGCGCATGTACTCCAGCCCTGGGTGGAAGAGCGAACTGTCTGTGAAAAAATAAAA	6257,7
QY	861,6	AAATTAATAATAATAATATG	863,6
Db	6257,8	AAAAACAAGATTCACCTTG	6259,8

```

RESULT 13
US-09-764-891-8396
; Sequence 8396, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8396
; LENGTH: 27154
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8396

Query Match          3.5%; Score 347.2; DB 11; Length 27154;
Best Local Similarity 58.6%; Pred. No. 4.5e-78;
Matches 899; Conservative 0; Mismatches 489; Indels 152; Gaps 11;

QY      7195 CAGGCCGGGCGTGGTGAATC-ACGCGCTTAATCTTAGCACTTTGGAGGCTGAGGCAAGT 72533
Db      542 CAGGCCGGGCGTGGTGAATC-ACGCGCTTAATCTTAGCACTTTGGAGGCTGAGGCAAGT 601
QY      7254 GGATTTGCTTGAAGCGGCTTGAGCCTTAGAGATTGAGACCAAGCGCTGGGCAATCGGCCAAAC 73133
Db      602 GGATTCACCTGAGGATC-----AGGAGTTCAAGAACCAAGCGCTTGAGCCTTGGTGAAC 652
QY      7314 CTCATCTCTTAACAAAATAATATAAATTAGTCGGGAGTGTAGTGCGTTCTATAGTCCCA 73733
Db      653 C-CATCTCTTAACAAAATAATATAAATTAGTCCAGGCGGTATATGAGCACTGTATATCCCA 711
QY      7374 TCTACTTCAGAGGCTGAG-----CCCGAAGGTCAGAGCTTCAGTGA 7415
Db      712 GCTACTCGGAGGCGTGAAGCAAGAAATCGCTTGAAACCAAGGAGCAAAAGTTGCAAGTGA 771
QY      7416 GCCGTGATCGTGCTACTGCACTCCAGCGCTGGGTGACAGATGAGACCATGTCTCAAAA-- 74733

```

Db	772	GCCAAAGTTCGTGCACCTGCACCTCCAGCTGGGTGACAAAGCAAGATCTCCATCTCAAAAAA	831
Oy	7474	-----AAAACAAAAA	7488
Db	832	TAATGTTGTTATTGTACATAAATTATACATAAATTCATGCACATAAATCTCAATTAAGTT	891
Oy	7489	AACAAAACAAAACAAACAAACAAACAAAACCACTATATATATATATATATATCTAGCTGA	7548
Db	892	TATTTAAGAACAAAAGTTAAACAGTAAATGAGATTCAACGGAGACATATGAATCTAGACAC	951
Oy	7549	GGTAGAATGCACATATTTGGTAAATATACCAACATGATACCAGCTACAGCATGGGGCAGT	7608
Db	952	TTGGAAAGAGAGATGATATGTTGATATTTGAATACAGGTAAATTTTGAACAAAATTTTCAATAC	10111
Oy	7609	CCCTCCCTCTCACTGGTAAATTTTCTCTGACTCAAGTTTGTG- TTGTGTT	7667
Db	1012	AATAAACATTCACATCTTTTATTAATTAATAAAAAAAGATGGAGCTACTGTAAATATTTTA	10711
Oy	7668	GCTGTGTTTGAAGATGAGATCTCACTGTGTACCCAGGCTGGAGTGCATAGGGCAATCT	7727
Db	1072	TTTTTTTTTTTGAAGCAAGATCTCACTGTGTGCTAGGCTGGAAATACAGTGGCAATGATCT	11311
Oy	7728	TGTTTCACTGCAACCTGTGCTCCTGGGTTCAAGCGATTCCTCCGCTCAGCCTCCGCTA	7768
Db	1132	CGGCTCACTGCAACCTGTGCTCCTCCGGGTTCAAAAAATCTCTGTGCTCAGCCTTCAGAG	11911
Oy	7788	TAGCTGGAGCTACAGCGCATATCCACCATGCTGGCTAAATTTTGTATTTTGTGGGT	7847
Db	1192	TAGCTGGAGCTACAGGTGCTCCACACACACACGCGCGC-----	1227
Oy	7848	TACATGTACTATTTATTAATTTATTTTGTATTTTATAGATAGATAGGTTTCAACAT	7907
Db	1228	-----TATTTTGTATTTTATAGTAGAGAGCGGGGTTTCAACAT	1265
Oy	7908	GTTGGCCAGGCTGTCTCGAATCTCTGTAACCTCAGGTATCCGCTGCTGGCCTCCCA	7967
Db	1266	GTTGGGCAAGGCTGTCTCGAATCTCTGTAACCTCAGGTATCTCCGCTCAACTTCCCA	1329
Oy	7968	AGTCTAGGATTTACAGGCGATGAGCAACAGCCTGGGCCCCCTCATAGGTTTATCTATTC	8027
Db	1326	AGTCTGGGATTTACAGGCGTGAAGCAGGCACTGGCTACTTGAATATTTACATATTT	1385
Oy	8028	TCATTGCTTCTCAACATTTGGCTTGCAGGTGACATCATATGTTCTCTCCACTTTCTCA	8087
Db	1386	AATTCACCT--TACAAAGGTGTATCATCATGTTTGTATTAATCAATAAAGGTTATCA	1443
Oy	8088	CTACTTCATGATCTTTCAAGTCTAGTTCCAACTGATATCTCCCTCAGTGTCTTTTTC	8147
Db	1444	ATTAATAATCTACTAGATGTTGAGCAATTTTTCACATTTATCAGCCAGTGTGATTTTTTTTT	1503
Oy	8148	CTAGTAAAGATTTCCAGAGAGGGAATCTGAATGCGCCAGTCCATATTTTACAGACACCA	8207
Db	1504	TTGAATGCTCTACTGTGTGCTTTTGGCAATTTTCTATTAATGATGTGTTTCTCTTTCAT	1563
Oy	8208	CATTAAAGTGTGATTGTCAGCCCTATGATATGG---CTACATTAATGGGTGGGAAT	8263
Db	1564	TTGTTTACAGGCCATCTGCTCATCAGCCTGGGAGTTTCCACAGACAAAGAACTGATTT	1623
Oy	8264	CATCATTTTACTTCATTTGCAAAAGCAGATAGCTGTGTTCT-----	8305
Db	1624	TCATTTTCTTCTCATCTCCAGGCTCATGATCAGTCTGTGATTTAGTTCATGAATGTTT	1683
Oy	8306	-----CAAAATAGGGCCCCGTGGGCCAGGTTGTTGGTCTCATGCTATATCCCAACATGT	8361
Db	1684	GTTAAATTTGAATGGAATGAGGCCAGGTGCAGTGGCTCAGGCTGTATATCCCACATTTT	1743
Oy	8362	GGAGAGCCGAGGGGGGAGATCATCTTGAATCTCAGAGTTCATGACAGCCTGGGCAACAT	8421
Db	1744	GGAAGCCCAAGGCTGTGTGATCATTAATGATGCAAGAAATTTGAAGACAGCCTGGCCAACAT	1803
Oy	8422	GGTGAATCTCATCTTCACTTAAATATCAAAAAATTAGCCAGGTGTGTGCTATGCAACA	8481

Db 1804 GGTGAAACCTATCTCTAATAAAATAC-AAAAATTAGCCGGCTGTGGCAAGCACT 1862
Qy 8482 GTATGCCAGCTGTTGAGAGGCTGAGGTGGAGATGCTCGAGTGTGGAGGACAGA 8541
Db 1863 GTATGCCAGCTGTTGAGAGGCTGAGGTGGAGATGCTCGAGTGTGGAGGACAGA 1922
Qy 8542 TTGAGTGAACCGTGAAGCTGCTCTGCAATCCAGCCTGGGTGACAGATTGAGACCTGT 8601
Db 1923 TTGAGTGAAGCTGAGATCGAGCACTGCACCTCCAGCCTGGGTGACAGATTGAGACCTGT 1982
Qy 8602 CTCAAAAAACAATTAATAATAATAATAT 8635
Db 1983 CTCAAAAAACAATTAATAATAATAATAT 2016

RESULT 14

US-10-034-650-10/c
; Sequence 10, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 53106
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-034-650-10

Query Match 3.5%; Score 346; DB 13; Length 53106;

Best Local Similarity 58.9%; Pred. No. 1,46-77;
Matches 850; Conservative 0; Mismatches 500; Indels 94; Gaps 11;

Qy 7192 ACCGAGCCGCGGCTGTGATGATCAGCTGTAACTCTGACACTTGGAGGCTGAGGAC 7251
Db 32946 ACATCGGCTGGGACGGTGGCTATGCTTAATCCAGCACTTGGAGGCTGAGGAC 32887
Qy 7252 GTGAGTGTCTTGAAGCGCTGAGCTTGAAGTTGAGACCAAGCTGGGCAATGSCAAA 7311
Db 32886 GTGAGTGTCTTGAAGCTTGAAGTTGAGACCAAGCTGGGCAATGATGAAA 32836
Qy 7312 ACCGATCTCTACAAAATAATAATAATAGTGGGTGTGTGTTCTATATCTCC 7371
Db 32835 ACCCGCTCTACAAAATAATAATAATAGTGGGTGTGTGTTCTATATCTCC 32776
Qy 7372 CATCTACTGAGAGCTGAGCCCGCA-----AGGTGAGGCTTCACT 7413
Db 32775 CAGCTACTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 32716
Qy 7414 GAGCGGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 7472
Db 32715 GAGCGGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 32656
Qy 7473 AAAAAACAAAACAA-----AAACAAACAAACAAACAAACAAACAAACAAACAAAC 7526
Db 32655 AAAAAACAAAACAAAGCTTGGATCTTAATACCTTAATGATGATGATGATGATGATGAT 32596
Qy 7527 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7586
Db 32595 GAGTTTGTGGGAGCACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 32536
Qy 7587 CCAAGCTACAGCATGGGGAGTCCCTCCCTCTCACTGATTAATTTTCTTCTCTGACT 7646

Db 32535 CTTAATCTCTTCAATTTCTAATAATAATAGATCCAGATTAATCCATTAATTTGTTCT 32476
Qy 7647 CACAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7706
Db 32475 TTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 32419
Qy 7707 TGAAGTGAATGAGCGCAATCTTGTGTTCACTGCAACCTCTGCTCTGGGTTCAAGCATC 7766
Db 32418 TGAAGTGAATGAGCGCAATCTTGTGTTCACTGCAACCTCTGCTCTGGGTTCAAGCATC 77359
Qy 7767 CTCTGCTGAGCTCTCCGATATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 7826
Db 32358 CTCTGCTGAGCTCTCCGATATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 32302
Qy 7827 TTTTGTATTTTTTGTGGGTACAATGATCTATTAATTAATTTTGTATTTT 7886
Db 32301 -----TAATTTTGTATTTT 32285

Qy 7887 GTAGAGTGAAGGTTTCAACCAATGTTGACAGGCTGTCTGCACTCTGACCTGAGTGTAT 7946
Db 32284 GTAGAGTGAAGGTTTCAACCAATGTTGACAGGCTGTCTGCACTCTGACCTGAGTGTAT 32227
Qy 7947 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8006
Db 32226 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 32167
Qy 8007 CTGATGCTTTTATCTATCTC-----TTGCTCTGCACTTGTGGCTGACGCTG 8061
Db 32166 CATCCATATATTTGAACCAACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 32107
Qy 8062 ACCATCATGTTCTCTGCACTTCTGCACTTCTGCACTTCTGCACTTCTGCACTTCTGCA 8121
Db 32106 GAACAAAGCAAAATTCAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 32047
Qy 8122 ATACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8181
Db 32046 AAAAAAACTGCTCTATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 31987
Qy 8182 CAGTCATATTTTCTGACACCACTGATTAATGCTGATTTGACCAAGCTATGTTG 8241
Db 31986 CATTAACCTTAAGCAAAATTCAGATTTATTTATTTATTTATTTATTTATTTATTTATTT 31927
Qy 8242 CTACATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8298
Db 31926 GTGATCATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 31867
Qy 8299 TGGTTCTCAAAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8358
Db 31866 ATCTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 31807
Qy 8359 TGTGGAGGCGGAGGCGGAGATCACTTGAATCCAGAGTTTGAACCAAGCTGAGGCA 8418
Db 31806 TGTGGAGGCGGAGGCGGAGATCACTTGAATCCAGAGTTTGAACCAAGCTGAGGCA 31747
Qy 8419 CATGTAATATCTCATCTCTCTAATAATAATAATAATAATAATAATAATAATAATAATA 8478
Db 31746 CATGTAATATCTCATCTCTCTAATAATAATAATAATAATAATAATAATAATAATAATA 31688
Qy 8479 CAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8538
Db 31687 TCTGTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31628
Qy 8539 AGATTCAGTGAACCGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8598
Db 31627 AGATTCAGTGAACCGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31568
Qy 8599 TGTCT 8602
Db 31567 TGTCT 31564

RESULT 15

US-10-364-505-9/c

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 10:53:38 ; Search time 12512.6 Seconds
(without alignments)
19424.068 Million cell updates/sec

Title: US-09-900-448-3_COPY_1_10000

Perfect score: 10000
Sequence: 1 tccctcccccagcagcgc.....agctgatttcgagcctaacc 10000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST.*
1: em_estda:*
2: em_estdm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	664.4	6.6	1201	9	AL534855 AL534855
2	550.4	5.5	552	2	BSM097187 BSM097187
3	493.2	4.9	566	10	BG002737 BG002737
4	436.4	4.4	497	28	AQ138990 AQ138990

Result No.	Score	Query Match	Length	ID	Description
5	424	4.2	424	2	BSM089286 BSM089286
6	410.8	4.1	430	2	BSM089310 BSM089310
7	331.4	3.3	537	2	BSM069237 BSM069237
8	287.6	2.9	333	10	BSM27475 BSM27475
9	286	2.9	340	12	BG954333 BG954333
10	272.2	2.7	6146	28	AQ839831 AQ839831
11	270	2.7	1201	9	AL534854 AL534854
12	268.8	2.7	1850	11	BC025771 BC025771
13	265.2	2.7	2286	11	BC035158 BC035158
14	255	2.5	1850	11	BC025771 BC025771
15	254	2.5	365	9	AV655253 AV655253
16	252.4	2.5	2296	11	BC035158 BC035158
17	251	2.5	985	11	BC035158 BC035158
18	250	2.5	1201	9	AL534301 AL534301
19	249.8	2.5	985	11	BC035158 BC035158
20	248.2	2.5	3828	11	BC035158 BC035158
21	245	2.5	1201	9	AL534301 AL534301
22	239	2.4	2551	11	BC024155 BC024155
23	236	2.4	1034	12	BM806909 BM806909
24	236	2.4	3777	28	AF101957 AF101957
25	235.8	2.4	1201	9	AL565216 AL565216
26	235.6	2.4	420	29	BZ59910 BZ59910
27	235.2	2.4	469	13	BU660220 BU660220
28	235	2.4	566	9	AL597545 AL597545
29	234.6	2.3	651	29	AG156348 AG156348
30	234	2.3	1103	12	BM904488 BM904488
31	233.6	2.3	378	14	CD516054 CD516054
32	233.4	2.3	340	28	AQ230483 AQ230483
33	233.4	2.3	505	28	AZ254926 AZ254926
34	233.4	2.3	609	29	AG149236 AG149236
35	233.4	2.3	873	14	CA454910 CA454910
36	232.8	2.3	1103	12	BM904488 BM904488
37	232.6	2.3	690	29	AG105245 AG105245
38	232.6	2.3	770	13	BX090291 BX090291
39	232.4	2.3	657	13	BU626531 BU626531
40	231.6	2.3	437	2	BSM082022 BSM082022
41	231.4	2.3	648	10	BE158921 BE158921
42	231.2	2.3	1080	13	BQ213854 BQ213854
43	231	2.3	2732	11	BC004387 BC004387
44	231	2.3	628	28	BS4637 BS4637
45	230.6	2.3	762	28	AQ311398 AQ311398
			493	9	AI733523 AI733523

ALIGNMENTS

RESULT 1
LOCUS AL534855
DEFINITION AL534855 Homo sapiens PITAL BRAIN Homo sapiens CDNA clone
ACCESSION AL534855
VERSION AL534855
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 13, 2001 this sequence version replaced gi:12798348.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5958.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODP006B110P1&cluster=5958.r. Contact :
Feng Liang Email: fliang@lifetech.com URL:

AL534855 1201 bp mRNA linear EST 12-MAY-2003
CSODP006Y122 5-PRIME, mRNA sequence.
AL534855
AL534855.2 GI:30541059
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polyes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12798348.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5958.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODP006B110P1&cluster=5958.r. Contact :
Feng Liang Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com/InvitrogenCorporation1600
 Faraday Avenue Genoscope Sequence ID : CS0DF006B11QPI.
 Location/Qualifiers

FEATURES
 source

1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF006Y122"
 /issue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_id="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
 BASE COUNT 250 a 330 c 303 g 278 t 40 others
 ORIGIN

Query Match 6.6%; Score 664.4; DB 9; Length 1201;
 Best Local Similarity 99.8%; Pred. No. 0.00025;
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5242 AGCAAGAGAGAGGCGTACCCGGTCTCACCAATGCTGTGATTGATGACAAAGT 5301
 DB 61 ATACAAAGAGAGGCGTACCCGGTCTCACCAATGCTGTGATTGATGACAAAGT 120
 QY 5302 GACCGGAGAGGCTTCTTGGGACTTGGCTACGGGAACATGAAAGAGCTTCTTGGCCAGCT 5361
 DB 121 GACCGGAGAGGCTTCTTGGGACTTGGCTACGGGAACATGAAAGAGCTTCTTGGCCAGCT 180
 QY 5362 GTTGGGAAGCTCTCTGCGCTGAGATGAGTGGGCGGCTACTAGCTTCCAGGGGTAC 5421
 DB 181 GTTGGGAAGCTCTCTGCGCTGAGATGAGTGGGCGGCTACTAGCTTCCAGGGGTAC 240
 QY 5422 CAATTCCTGCTTCAACCTCTGTCAGGGAGAGGTGCTCCAGAGTACCCGCGGATGTC 5481
 DB 241 CAATTCCTGCTTCAACCTCTGTCAGGGAGAGGTGCTCCAGAGTACCCGCGGATGTC 300
 QY 5482 CGAGACTACTTCAATGCTCTGCTGCGGAGAGTGAAGAACCTTGAAGACTTGT 5541
 DB 301 CGAGACTACTTCAATGCTCTGCTGCGGAGAGTGAAGAACCTTGAAGACTTGT 360
 QY 5542 CAGATTCAATCACTTCCCTGAGCTGTGATGATGATGATGATGATGATGATGATGATGAT 5601
 DB 361 CAGATTCAATCACTTCCCTGAGCTGTGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 5602 TCCGTGTGGAGACCTTGGCCCTTAATCTAGCCCACTTCAATCTGATTTTCCCAT 5661
 DB 421 TCCGTGTGGAGACCTTGGCCCTTAATCTAGCCCACTTCAATCTGATTTTCCCAT 480
 QY 5662 GCGCTCAATGAGGGAACCCACACCCCACTAACCAGCCCACTTCTTCCACTTGAACCT 5721
 DB 481 GCGCTCAATGAGGGAACCCACACCCCACTAACCAGCCCACTTCTTCCACTTGAACCT 540
 QY 5722 CACTGTGACCTTGGCCCTTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 5781
 DB 541 CACTGTGACCTTGGCCCTTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 5782 ACAAGGAATGGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5841
 DB 601 ACAAGGAATGGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
 QY 5842 TAGCCCAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5901
 DB 661 TAGCCCAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
 QY 5902 CAGTGG 5907
 DB 721 CAGTGG 726

RESULT 2

HSN097187
 ID HSN097187 standard; RNA; EST; 552 BP.
 XX AC BX503470;
 XX SV BX503470.1
 XX DT 09-MAY-2003 (Rel. 75, Created)
 XX DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)
 XX DE Homo sapiens mRNA; EST DKFZp779H149_r1 (from clone DKFZp779H149)
 XX KW EST; expressed sequence tag.
 XX OS Homo sapiens (human)
 XX OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 XX OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 XX [1]
 XX RP 1-552
 XX RA Bioecker H., Boecker M., Mewes H.W., Weil B., Anld C., Osanger A., Robo G.,
 XX RA Han M., Wiemann S.;
 XX RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
 XX RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
 XX CC This is the 5' sequence of the clone insert.
 XX CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 XX CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 XX CC sequenced by GBR (National Research Centre for Biotechnology
 XX CC Ltd., Braunschweig/Germany) within the cDNA sequencing
 XX CC consortium of the German Genome Project.
 XX CC No BL sequence available.
 XX CC This clone (DKFZp779H149) is available at the RZPD in Berlin.
 XX CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
 XX CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
 XX FH Key Location/Qualifiers
 FH FT source 1..552
 FT /db_xref="taxon:9606"
 FT /mol_type="mRNA"
 FT /organism="Homo sapiens"
 FT /clone="DKFZp779H149"
 FT /clone_id="779" (synonym: hnccl). Vector pSPORT1_SfiI; host
 FT DH10B; sites SfiIA + SfiIB
 FT /dev_stage="fetal"
 FT /issue_type="liver"
 XX SQ Sequence 552 BP; 119 A; 163 C; 93 G; 177 T; 0 other;
 Query Match 5.5%; Score 550.4; DB 2; Length 552;
 Best Local Similarity 99.8%; Pred. No. 0.039;
 Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3366 GGGGATCAATCCAGCTTCCATATATACCTCCCTGAGTGTCAACCATCCTTGGC 3425
 DB 1 GGGGATCAATCCAGCTTCCATATATACCTCCCTGAGTGTCAACCATCCTTGGC 60
 QY 3426 TCTGAGGAATATCTATGCTGTGAGCACTGATCTAGATCTACTTCTGAGCTCAGACAT 3485
 DB 61 TCTGAGGAATATCTATGCTGTGAGCACTGATCTAGATCTACTTCTGAGCTCAGACAT 120
 QY 3486 CTCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3545
 DB 121 CTCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 3546 GATCTTCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3605
 DB 181 GATCTTCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 3606 AATGCTGATTTACTACAAACCAAGTGACACAGGATCTGATGATGATGATGATGATGATGAT 3665

Db 241 AATGTCCTGATTAACAAACCAAGTGACACAGGGCTCTGATGATCTGGCCCTGGTTA 300
 QY 3666 TCTTCAGGTTTATCTCTCCCGCCACATTCATCTGTGCGACGCCATAGAAATCTAC 3725
 Db 301 TCTTCAGGTTTATCTCTCCCGCCACATTCATCTGTGCGACGCCATAGAAATCTAC 360
 QY 3726 ATGAGGTTGAGACACACTGCTCTCTGATGTTGGGCTGACAGTGCCTCCCTGCTGG 3785
 Db 361 ATGAGGTTGAGACACACTGCTCTCTGATGTTGGGCTGACAGTGCCTCCCTGCTGG 420
 QY 3786 TAAACACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3845
 Db 421 TAAACACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 QY 3846 CCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3905
 Db 481 CCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
 QY 3906 TGTCTATATCT 3917
 Db 541 TGTCTATATCT 552

RESULT 3
 BG002737 566 bp mRNA linear EST 24-JAN-2001
 LOCUS BG002737
 DEFINITION QV4-GN0250-171100-549-d12 GN0250 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG002737 GI:12442358
 VERSION BG002737.1 GI:12442358
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 566)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,
 Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, P. F.,
 Goldman, G. H., Carvalho, A. F., Matukuma, A., Baia, G. S., Simpson, D. H.,
 Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,
 M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and
 Simpson, A. J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL 20202663
 MEDLINE 10737800
 PUBMED
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?fl=QV4-GN0250-
 171100-549-d12&t3=2000-11-17&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 566.

FEATURES

source
 1..566
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_id="GN0250"
 /notes="Organ: placenta normal; Vector: puc18; Site_1: SmaI
 ; Site_2: SmaI; A mini-library was made by cloning
 products derived from ORESTS PCR (U.S. Letters Patent
 application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the puc 18 vector. Reverse

BASE COUNT 127 a 147 c 123 g 169 t
 ORIGIN
 Query Match 4.9%; Score 493.2; DB 10; Length 566;
 Best Local Similarity 97.8%; Pred. No. 0.37;
 Matches 532; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

QY 6209 CCTGAGCTCTCCCTGCTCCCTGACCTCTATGTCCTGTCTTACTCTTGGCTCCCTT 6268
 Db 26 CCTGAGCTCTCCCTGCTCCCTGACCTCTATGTCCTGTCTTACTCTTGGCTCCCTT 82
 QY 6269 TGTCTGATATCTCTTCTGAGTCTCTGAGCTCCCTGATTTATCTCTGAACTCATC 6328
 Db 83 TGTCTGATATCTCTTCTGAGTCTCTGAGCTCCCTGATTTATCTCTGAACTCATC 142
 QY 6329 TTGTTTCAAGTTCCTGTTCTCTATGTCAGACCCCTGGGCATAGACCTGCGGATGA 6388
 Db 143 TTGTTTCAAGTTCCTGTTCTCTATGTCAGACCCCTGGGCATAGACCTGCGGATGA 202
 QY 6389 GATGTTCTCATGTCGAGAACCACTGAGAGTGTGGTACTTTAGACCTTTAGAGCT 6448
 Db 203 GATGTTCTCATGTCGAGAACCACTGAGAGTGTGGTACTTTAGACCTTTAGAGGCT 262
 QY 6449 GGCTTCACAGCTCTGAGAGTGTTCCTCTGAGTGAAGCAAGAGATACCCCTCCCTG 6508
 Db 263 GGCTTCACAGCTCTGAGAGTGTTCCTCTGAGTGAAGCAAGAGATACCCCTCCCTG 322
 QY 6509 ACCCTGGCATCAATTTGTTAAAGCATCTAATATACCTAGAGGCTGTTGATTCAGT 6568
 Db 323 ACCCTGGCATCAATTTGTTAAAGCATCTAATATACCTAGAGGCTGTTGATTCAGT 382
 QY 6569 CAGGAGTAATATGTCATGTCGACAGTGTGAGATATCCCAAGAGAGTGAACACAC 6628
 Db 383 CAGGAGTAATATGTCATGTCGACAGTGTGAGATATCCCAAGAGAGTGAACACAC 442
 QY 6629 ATGACATCCACCTGAGATATATATATATGAGACAGTGTGAATATATATATATCTGA 6688
 Db 443 ATGACATCCACCTGAGATATATATATATGAGACAGTGTGAATATATATATATCTGA 502
 QY 6689 AATATAT-TTTTTCACACAAATTTTGGCTCTGACCTCTGACAAATTTGACAGT 6747
 Db 503 AATATCTCTTTTTCACACAAATTTTGGCTCTGACCTCTGACAAATTTGACAGT 562
 QY 6748 TATG 6751
 Db 563 TATG 566

RESULT 4
 AQ138990 497 bp DNA linear GSS 24-SEP-1998
 LOCUS AQ138990
 DEFINITION HS 3088 B2 C01 MR CIT Approved Human Genomic Sperm Library D Homo
 sapiens genomic clone plate=3088 Col-2 Row=F, genomic survey
 sequence.
 ACCESSION AQ138990
 VERSION AQ138990.1 GI:3529643
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 497)
 Mahairas, G. G., Wallace, J. C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M. D. and
 Hood, L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 JOURNAL 99380589
 MEDLINE 10449764
 PUBMED
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L

Df		64	CAGCATTCTTCACCCCTTGAGACTGCACTTGGCGTTCTGTTCCTGC	5
Oy		5757	ACCCT 5760	
Rt		4	Accc 1	
Df				
RESULT 6				
ID	HSN089310/C	standard; RNA; EST; 430 BP.		
XX	HSN089310			
XX	BK496085;			
SV	BK496085.1			
DT	09-MAY-2003 (Rel. 75, Created)			
DT	09-MAY-2003 (Rel. 75, Last updated, Version 1)			
DE	Homo sapiens mRNA; EST DKFZp779C1326_r1 (from clone DKFZp779C1326)			
XX	EST; expressed sequence tag.			
KX				
OS	Homo sapiens (human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia			
OC	Eutheria; Primates; Catarrhini; Homnidae; Homo.			
XX	[1]			
RA	1-430			
RP	Bloecher H., Boecher W., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.			
RA	Han W., Wiemann S.;			
RL	Submitted (07-MAY-2003) to the EMBL/GenBank/DDBJ databases.			
MPS,	Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY			
CC	This is the 5' sequence of the clone insert			
CC	clone from S. Wiemann, Molecular Genome Analysis, German Cancer			
CC	Research Center (DKRZ); Email s.wiemann@dkrz-heidelberg.de;			
CC	sequenced by GBF (National Research Centre for Biotechnology			
CC	Ltd., Braunschweig/Germany) within the CDNA sequencing			
CC	consortium of the German Genome Project.			
CC	No sl sequence available.			
CC	This clone (DKFZp779C1326) is available at the RZPD in Berlin.			
CC	Please contact the RZPD: Ressourcentum, Heubnerweg 6,			
CC	14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzdp.de			
FX				
FH	Key Location/Qualifiers			
FT				
FT	source			
FT	1..430			
FT	/db_xref= "taxon:9606"			
FT	/mol_type= "mRNA"			
FT	/organism= "Homo sapiens"			
FT	/clone= "DKFZp779C1326"			
FT	/clone_lbb= "779 (synonym: hnccl). Vector pSpartil_Sfi; host			
FT	DHI0B; sites SfiI + SfiIB"			
FT	/dev_scage= "fetal"			
FT	/tissue_type= "liver"			
SQ	Sequence 430 BP; 111 A; 97 C; 142 G; 80 T; 0 other;			
Oy	Query Match	4.1%; Score 410.8; DB 2; Length 430;		
Df	Beet Local Similarity	97.2%; Pred. No. 12;		
Df	Matches 418; Conservative 0; Mismatches 12; Indels 0; Gaps 0;			
Oy	5331 GGGAACCATGAAGAAGCGTTCCTGAGCAGCTGTTGGAACTGCTCGGCCGATGATG	5390		
Df	430 GGGAACCATGAAGAAGCGTTCCTGAGCAGCTGTTGGAACTGCTCGGCCGATGATG	371		
Oy	5391 GCTGGCGCGCTACTACTGCTTCAGGGTAACAATTCTGGCTTCGACCTGTCAAGGG	5450		
Df	370 GCTGGCGCGCTACTACTGCTTCAGGGTAACAATTCTGGCTTCGACCTGTCAAGGG	311		
Oy	5451 AGAGTGCTCCCAAGTACCCCGGGAGATGTCCGAGACTATTAGCTCGCTGGCAG	5510		

```

Db      310 AGAGGTGCTCCACAGGATGACCCGGGGATGATGCCAGACTACTTTCATGCTCCCTGGCCCTGGCAG 251
Qy      5511 AGGTGAGAAAGACCCCTAGCACTTGTAGACCTGTGCAGATTTCATCACTTCCCTGAGCTGTG 5570
Db      250 AGGTGAGAAAGACCATGACACTTGTAGACCTGTGCAGAAATTCACTTCCCTGAGCTGTG 191
Qy      5571 GGATTCACAGTGTCTAGCTCTCACTTCTAACTCCGATGGCGACACTTGTAGCCCTTAATC 5630
Db      190 GGATTCACAGTGTCTAGCTCTCACTTCTAACTCCGATGGCGACACTTGTAGCCCTTAAAT 131
Qy      5631 TAGCCCATTTCCATTTCTGTGATTTTCCATTGCCCTCATATGGGAGAAACCCACACCCAC 5690
Db      130 TAGCCCATTTCCATTTCTGTGATTTTCCATTGCCCTCATATGGGAGAAACCCACACTCCAC 71
Qy      5691 TAAACCCAGGACATCTCTTCCACTTGTAGACTCACTGACCTTGTGCTCTCTGTGT 5750
Db      70 TAAACCCAGGACATCTCTTCCACTTGTAGACTCACTGACCTTGTGCTCTCTGTGT 11
Qy      5751 CTCCTCACCC 5760
Db      10 CTCCTCACCC 1
RESULT 7
HSM069237
ID      HSM069237 standard; RNA; EST; 537 BP.
XX      AC      BX479451;
XX      SV      BX479451.1
XX      09-MAY-2003 (Rel. 75, Created)
DT      09-MAY-2003 (Rel. 75, Last updated, Version 1)
XX      Homo sapiens mRNA; EST DKFPZ686C16211_r1 (from clone DKFPZ686C16211)
DE      EST; expressed sequence tag.
KM      Homo sapiens (human)
OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia
OC      Eutheria; Primates; Catarrhini; Homnidae; Homo.
XX      [1]
RN      1-537
RP      Koehler K., Beyer A., Mewes H.W., Well B., Amid C., Osanger A., Fobo G.,
RA      Han W., Wiemann S.;
RT      Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL      MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY
XX      This is the 5' sequence of the clone insert
CC      Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC      Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC      sequenced by BMFZ (Biomedical Research Center at the
CC      Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
CC      sequencing consortium of the German Genome Project.
CC      No s1 sequence available.
CC      This clone (DKFPZ686C16211) is available at the RZPD in Berlin.
CC      Please contact the RZPD: Resourcenzentrum, Heubnerweg 6,
CC      14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX      Key      Location/Qualifiers
FH      1. 537
FT      /db_xref="taxon:9606"
FT      /mol_type="mRNA"
FT      /organism="Homo sapiens"
FT      /clone="DKFPZ686C16211"
FT      /clone_1db="686 (synonym: hlc03). Vector pSport1_Sfi1 host
FT      DH10B; sites Sfi1A + Sfi1B"
FT      /dev_stage="adult"
FT      /issue_type="cDNA-collection"
XX

```

SQ Sequence 537 BP, 142 A; 110 C; 169 G; 116 T; 0 other;

Query Match 3.3%; Score 331.4; DB 2; Length 537;

Best Local Similarity 99.7%; Pred. No. 2.2e+02; Indels 0; Gaps 0;

Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCCCTCTCCCGACGAGCCGACGAAATCTGATGATTCAGACAGGCTTCTGACAGCTG 60

205 TCCCTCTCCCGACGAGCCGACGAAATCTGATGATTCAGACAGGCTTCTGACAGCTG 264

61 AAGCAAGTTGTTGAGGAATTCCTGATGAGATCATGGGCTCTCAGAGGAGATA 120

265 AAGCAAGTTGTTGAGGAATTCCTGATGAGATCATGGGCTCTCAGAGGAGATA 324

121 TAAGTTTCAGAGGCTGAGAGGAGAAAGATGAGGGGAGCTTAGAATGTGCTC 180

325 TAAGTTTCAGAGGCTGAGAGGAGAAAGATGAGGGGAGCTTAGAATGTGCTC 384

181 CCATTGCCCCAACCCAGAAAGAGATGCGCTGCAATGGGAGAGTGATGAGA 240

385 CCATTGCCCCAACCCAGAAAGAGATGCGCTGCAATGGGAGAGTGATGAGA 444

241 CATTGCTGTAGCAGCGATGCGATTCGCCAGGCTGCCAGAGACTCAGAGAGTCCAGCTT 300

445 CATTGCTGTAGCAGCGATGCGATTCGCCAGGCTGCCAGAGACTCAGAGAGTCCAGCTT 504

301 GCCCACTGACCTATGAGAGGAGATGATTTCA 333

505 GCCCACTGACCTATGAGAGGAGATGATTTCA 537

RESULT 8
BF827475 393 bp mRNA linear EST 13-JAN-2001

LOCUS CM4-HN0022-251100-461-h07 HN0022 Homo sapiens cDNA, mRNA sequence.

DEFINITION BF827475 GI:12171494

ACCESSION BF827475.1

VERSION BF827475.1

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 393)

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,

Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,

Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H.,

Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare

M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and

Simpson, A. J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT 10737800

CONTACT: Simpson A.J.G.

LABORATORY: Laboratory of Cancer Genetics

RUA Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="HN0022"

/note="Organ: head, normal. Vector: puc18; Site 1: Smal;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORFESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

[illegible]

Oy	7690	CACCTGTGCAACCAGGCTGGAGTGAATGGGGCAATCTTGTTCACTGCACACTCTGCT	7749
Dd	1272	CACCTCAGTAGCCCAAGCTGGAGTGAAGTGGAGTAATTTTGGCTCACTGCCAACTTCACCT	1668
Oy	7750	CTTGCGATTCAAGCGAATCCTCTGCTGCCTCACGCTCCCGTATAGCTGGACA-CT-AcAGCGCAT	7808
Dd	1667	CCCAGGTTCAGACAAATTCATAGCTCAGCCTCCGCTCGAGTACGTGGTATTAACAGGAGACG	1608
Oy	7809	ACCAACCATGCTGGCTGAATTTTGTATTTTTTTGGGTACAAATGTAATAATTATTAAT	7868
Dd	1607	ACTACCAAGCTGGC-----	1593
Oy	7869	TTAATTTTGTATTTTGTAGAGAATAGGTTTCAACATGTTGGCAGGCTGTCTCGAA	7928
Dd	1592	-TAACTTTTGTATTTTGTAGAGAGACAGGTTTCACTATGTTGGCAGGCTGTGATCAA	1534
Oy	7929	CTCTGACCTCAGGTATCGCGCTGCTCGGCTCCGACCTCCAAAGTGTAGATTACAGCAATG	7988
Dd	1533	CTTTGAACCTCAGAGTATCTCGCGCTCGGCTCCGAAAGTGTGGATCACAGTGGG	1474
Oy	7989	AGCAACCAAGCTGTGGCCCCCTCATAGSTTTTATCTATCT-----CTTGCT	8035
Dd	1473	TGCATCATGCGCCGCGCTTCACACATTCATCAAAATCTGCCCTTCTCTCCACTTCTCT	1414
Oy	8036	TCTTCACAACCTTGGCTTGCACGTCGACATCATATTC-----TC	8075
Dd	1413	ATTCTCTGTATTTGGAGACGTTAGTAATACCAATAGATTCAGTGGAAAACAGAAACCTTT	1354
Oy	8076	TCCAATTTTCATCTATCTCATGATCTTTCAGTCTCAGTTCCAATGATATCTCTCTCAGT	8135
Dd	1353	GGCCTTATATCTCCCTCTCTCCTCTTATTTATGTTAGCAATCTTACCATTCTACCTTCATA	1294
Oy	8136	TGCTCTTTTCTCAGTAAGATTTCCAGAGAGGAATGSAATGSGCCAGTCAMATTTT	8195
Dd	1293	TCTTATACCAATCCTTATATTTTCACTTCCACTGCCAACATAGCCCTCTATCTCGGG	1234
Oy	8196	CAGACCA-----CCACATTAAGTGTGATTGCGAGCCTATGTAATGGCTACATTA	8249
Dd	1233	CTATCAGAAATACTTCTACACATGATTTTCTATTAGCATCATCTTCAACATATCTGTT	1174
Oy	8250	ATGGGTTGGGAATCATCATTTACTTCACTTGCACAAAGACATBACTCTGTTCTCAA	8309
Dd	1173	CTCTGTATTTCAATTAATCTCTTGCAGCCTGTCTTA--TTCTGTCAATAGCCCTCTCAA	1115
Oy	8310	ATAAGGCCCCGTGGGCGAGGTGAGTGGCTCATAGCTTAATCCCAACCTGAGGAGGCC	8369
Dd	1114	AATCCCCATACCGGCTGGGCGGGGATCAACGGCTGTAAATCCAGCATTTTGGAGGCC	1055
Oy	8370	GAGGGGGGAGATCATTTGAGTCCAGAGTTCTAGACCAGCCTGGGCAACATGTGTAAT	8429
Dd	1054	GAGGTGGCGGATCATGTAAGTGGGAGTTTGAGAACCAAGCTGTACCAATGAGGAAC	995
Oy	8430	CTCATCTTCAATAAAAAATACAAAAAATTAAGCAGGTGTGTGGCATGCAACCATATGTC	8489
Dd	994	CCCGCTCTTACAAAAAATAC--AAATTAAGCGGGCGTGGCGGCATGCTGTAAATCC	937
Oy	8490	AGCTGTTCAAGAGGCTGAGGTGGAGATGCTCAGTGTGGAGGCGAGATTGCAAGT	8549
Dd	936	AGCTACTGGAGACTGAGCGAGGAATTCGTTTBAACCCGGAGGCGGAGGTTGCGGTG	877
Oy	8550	AACGTTACTGTGCTCTGCAATCCAGCCTGGGTAC-AGATTGACCTGTCTCAAA	8608
Dd	876	AGCGAGATGGCGCATTTGCACTCCAGCTGGGCAACAAAGACAAAATCTGTGCTCAAA	817
Oy	8609	AACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	8632
Dd	816	GAGAAACACTACAAACAAAAAA 793	

RESULT 13			
LOCUS	BC035158	2296 bp	mRNA
DEFINITION	Homo sapiens, clone IMAGE:5265332, mRNA.		
			linear
			HTC 04-MAR-2003

VERSION	BC035158
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 2296)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Miklos Palokovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shireki Tohliyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbioology.org contact: amadan@systemsbioology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Ketreman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: http://image.llnl.gov Series: IRAC Plate: 73 Row: 1 Column: 12 This clone has the following problem: retained intron.</p> <p>Location/Qualifiers 1..2296</p> <p>/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5255332" /clone_type="Testic" /gene_id="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescript"</p>
BASE COUNT	623 a 575 c 568 g 530 t
ORIGIN	
Query Match	2.7%; Score 265.2; DB 11; Length 2296;
Best Local Similarity	60.0%; Pred. No. 1e+03;
Matches 608; Conservative	0; Mismatches 308; Indels 98; Gaps 6;
QY	7657 TTGTTGTTGTTGCTGTTGTTGTTGAGATGAGTCTACTCTGTCAACCGAGGCTGAGAGTCA 7716
DB	1298 TTTCTCTTTTTTTTTTTTTTTCAGACAGAGCTCTACTCTGTGCACCCGGCTGAGTCAAG 1357
QY	7717 TGGGCAATCTGGTTGACAGCAACCTGTGCTCTGGGTTCAAGCATCTCTGAGCTC 7776
DB	1358 TGGCAATCTCGGCTCACTGCAACCTGTGCTCTCCAGTTCAAGTATCTCTGCTC 1417
QY	7777 AGCTCCCGTATGACTGGAAGCTAGAGGCGCATACACATGCTGAGTATTTTGTATT 7836
DB	1418 AGCTCCCGGTACTGAGATTAGAGGACGTGCAACAGCCCTGGC----- 1464
QY	7837 TTTTGTGGTTACATGTAATTAATTAATTTTGTATTTTGTATTTAGTAGAGATAG 7896
DB	1465 -----TAAATTTTGTATTTTGTATTAAGATAG 1491
QY	7897 GGTTCACAAATGTTGGCAGGCTGTCTGCAATCTGCACTCGAGGTAGTACCGCTGCT 7956
DB	1492 GGTTCACAAATGTTGGCAGGCTGTATCTTGAATCTGCACTCAAGTATTTGCTGCT 1551
QY	7957 CGGCTCCCAAGTGTAGATTACAGAGCAATGAGCAACAAGCTGAGCCCTCATAGATT 8016
DB	1552 CAGCTCCCAAGGCGCTGTATTAAGGCGTGAATACCGTGGCCCAAGCCCTGTTAGACT 1611
QY	8017 TTATATATATCTTT--TGCTTTTCAAACTTTGGCTTGACGTGACATATATGTTCT 8074


```

Db      1349 AACCCAAAGGTTCTGTTTCAACTGATCTATTGTATTAATACTCAATACCA 1408
Qy      8274 TTGATTCGCAAAAGCAGCATAGCTCTGTCTCAAAATAGGGCCCTCGGCGAGGTGG 8333
Db      1409 GGAATAGGAAAGTGGAGGAGGAGCATTTGATAGATGTGTAGAGGACCCGGCATGGA 1468
Qy      8334 TGGCTATGCTTAATATCCCAACACTGTGGAGGCGGAGGAGGAGATCACTTGAAGTCC 8393
Db      1469 TGGCACCACCTGTGTATCCAGCATTTGGAGAGCCGAGGCGAGATCACTCGAGTTC 1528
Qy      8394 AGGAGTTCTAGACACAGCTGCGGCAATGATGAAATCTCATCTAATAAAATACAAA 8453
Db      1529 AAGAGTTGATACAGCTGCGGCAATAGTGAACCTGTCTCTAATAAAATACAAA 1588
Qy      8454 AATTAGCCAGGTGTGTGGATGACCAAGTAGTCCAGCTGTTCAGAGGCTGAGTGG 8513
Db      1589 GTTAGCCAGGCGTGTAGTGTGTCCTCTGTTAATACAGCTACTCAGAGGCTGAGGATG 1648
Qy      8514 AGGATTCCTGAGTGTGGAGGAGGAGATTCAGTGAACCGTACCTGTGCTGCAATC 8573
Db      1649 AGAATGCTTGAACCTGGAGGTGAGAGTTCAGTGAACCAATTAAGCCTGCACTC 1708
Qy      8574 CAGCTGGGTGACAGATTGAGACCTGTCTCAAAAAACAATTAATTAATTAAT 8633
Db      1709 CAGCTGGGCTAGTGAAGTGAAGTCTGTCTCAAAAAAATTAATTAATTAAT 1768
Qy      8634 A 8634
Db      1769 A 1769

```

RESULT 15

```

LOCUS      AV655253          365 bp      mRNA      linear      EST 15-JAN-2002
DEFINITION AV655253      GLC Homo sapiens cDNA clone GICERD03 3', mRNA sequence.
ACCESSION  AV655253
VERSION     AV655253.1      GI:9876267
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE

```

AUTHORS    Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.

```

TITLE

```

Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

```

```

JOURNAL    MEDLINE
PUBMED     11752456

```

COMMENT

```

Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@hgsc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1..365

```

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GICERD03"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_id="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:

```

```

BASE COUNT      69 a      126 c      83 g      87 t
ORIGIN
Query Match      2.5%; Score 254; DB 9; Length 365;
Best Local Similarity 100.0%; Pred. No. 6.3e+03;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      5654 TTCCCATTCGCCCTGCAATATGGGGAACCAACCCCACTAAACCCAGCATCTTTCCACC 5713
Db      1 TTCCCATTCGCCCTGCAATATGGGGAACCAACCCCACTAAACCCAGCATCTTTCCACC 60
Qy      5714 TTGGAACCTGACTGTGACTCTGGCCCTCTTGTGTCTCCCTCAACCATTTCTCTCCA 5773
Db      61 TTGGAACCTGACTGTGACTCTGGCCCTCTTGTGTCTCCCTCAACCATTTCTCTCCA 120
Qy      5774 GGCCATGACACAGGAATGGAGCTGGCCATGGGAACAGTACCAACCATGGCCCTGAGTAT 5833
Db      121 GGCCATGACACAGGAATGGAGCTGGCCATGGGAACAGTACCAACCATGGCCCTGAGTAT 180
Qy      5834 ATGCGCTGTAGCCCAATCTAGTCTTGTCTGCACTGACGTCTGACAAACCATGGCCACC 5893
Db      181 ATGCGCTGTAGCCCAATCTAGTCTTGTCTGCACTGACGTCTGACAAACCATGGCCACC 240
Qy      5894 TATGCTTCAGTGG 5907
Db      241 TATGCTTCAGTGG 254

```

```

Search completed: December 16, 2003, 05:34:16
Job time : 12524.2 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 06:39:26 / Search time 8787.02 Seconds

(without alignments) 17398.312 Million cell updates/sec

Title: US-09-900-448-3_COPY_10001_13737

Perfect score: 3737

Sequence: 1 cccaagagagagagtcagtcac.....accttttgcgtgtcttca 3737

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapept 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```

Database :
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sbs:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	3730.6	99.8	191556	9 AC068733	AC068733 Homo sapi
C 2	3729	99.8	112295	9 AC084337	AC084337 Homo sapi
C 3	2449	65.5	148198	2 AC024232	AC024232 Homo sapi
4	618	16.5	629	9 HUMHXXMA07	M36803 Human hemop
5	573.8	15.4	577	9 HUMHXXMA06	M36802 Human hemop
6	567.4	15.2	249769	2 AC125227	AC125227 Mus muscu
7	557.8	14.9	1467	10 NMHMPX8	X56830 M.musculus
8	520	13.9	238748	2 AC120746	AC120746 Rattus no
C 9	520	13.9	275982	2 AC131626	AC131626 Rattus no
10	444.4	11.9	1621	9 BC005395	BC005395 Homo sapi
11	443.2	11.9	1546	9 HUMHXXMA	X03048 Huma hemope
12	431.2	11.5	1485	9 HSHPEPR	X02537 Human mRNA
C 13	365	9.8	400	11 G11396	G11396 SHGC-10293
C 14	265.2	7.1	504	6 AX410520	AX410520 Sequence
15	255.6	6.8	1486	4 OCHEMOEX	X16429 O.cuniculus
16	206.4	5.5	1475	10 BC011246	BC011246 Mus muscu
17	205.6	5.5	1470	10 BC019901	BC019901 Mus muscu
C 18	204	5.5	339	6 AX523973	AX523973 Sequence
C 19	204	5.5	339	6 AX522711	AX522711 Sequence
20	199.6	5.3	1528	4 SSU14751	U14751 Sus scrofa
21	198.6	5.3	1516	10 RATHMPX	M62642 Rat (clone
22	191.6	5.1	1430	10 NMU89889	U89889 Mus musculu
23	160	4.3	160	11 G13507	G13507 human STS S
24	136	3.6	631	6 AX525553	AX525553 Sequence
25	136	3.6	689	6 AX526031	AX526031 Sequence
26	116.2	3.1	107527	9 AC005018	AC005018 Homo sapi
C 27	116.2	3.1	186349	2 AC073420	AC073420 Homo sapi
28	112.8	3.0	60036	2 AC009610	AC009610 Homo sapi
C 29	112.8	3.0	60036	2 AC009610	AC009610 Homo sapi
C 30	112.8	3.0	132003	9 AC138779	AC138779 Homo sapi
C 31	112.8	3.0	163404	2 AC009614	AC009614 Homo sapi
C 32	112.8	3.0	164084	2 AC026580	AC026580 Homo sapi
C 33	112	3.0	1883	9 AK094459	AK094459 Homo sapi
C 34	112	3.0	93724	2 AC008451	AC008451 Homo sapi
35	112	3.0	109884	2 AC010481	AC010481 Homo sapi
36	112	3.0	120041	2 AC009038	AC009038 Homo sapi
C 37	112	3.0	178056	9 AC008536	AC008536 Homo sapi
C 38	112	3.0	181926	2 AC008461	AC008461 Homo sapi
C 39	109.2	2.9	92345	9 AC128687	AC128687 Homo sapi
40	109.2	2.9	154936	2 AC021252	AC021252 Homo sapi
41	109.2	2.9	182262	2 AC023963	AC023963 Homo sapi
42	109	2.9	104818	2 AC026690	AC026690 Homo sapi
43	109	2.9	159853	9 AC010451	AC010451 Homo sapi
44	108.6	2.9	195986	2 AC116933	AC116933 Papio anu
C 45	108.2	2.9	56631	9 HS702M17	AL034427 Human DNA

ALIGNMENTS

```

RESULT 1
AC068733/c 191656 bp DNA linear PRI 24-MAY-2002
LOCUS AC068733 Homo sapiens chromosome 11, clone RP11-304C12, complete sequence.
DEFINITION AC068733
ACCESSION AC068733
VERSION AC068733.12 GI:19703148
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 191656)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-304C12
JOURNAL Unpublished

```

REFERENCE
AUTHORS

2 (Bases 1 to 191656)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Buxte,G.,
Campopiano,A., Caselle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagoes,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Lander,T., Lehocsky,J.,
Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPeeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,T., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Direct Submission
Submitted (08-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (Bases 1 to 191656)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagoes,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Laroque,K., Lamazares,R.,
Lander,T., Lehocsky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,T., Peterson,K., Phunhhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Direct Submission
Submitted (27-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

4 (Bases 1 to 191656)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagoes,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
Lamazares,R., Lander,T., Lehocsky,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T., Peterson,K., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

COMMENT

Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 25, 2002 this sequence version replaced g1:19683105.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10266
Center clone name: 304_C12

Location/Qualifiers
1. 191656
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-304C12"
/clone_11b="RP11-11 Human Male BAC"
complement(189..1091)
/rpt_family="L2"
/rpt_family="GN)n"
complement(1151..1401)
/rpt_family="L2"
complement(1891..2187)
/rpt_family="AluUb"
3141..3165
/rpt_family="(CA)n"
complement(3205..3522)
/rpt_family="L2"
complement(4393..4591)
/rpt_family="L2"
4624..4932
/rpt_family="AluSp"
4987..5077
/rpt_family="MIR"
complement(5165..6245)
/rpt_family="Tigret3b"
complement(6246..6555)
/rpt_family="AluUb"
complement(6556..6634)
6940..7136
/rpt_family="L2"
complement(7607..7849)
/rpt_family="L2"
complement(11242..11318)
/rpt_family="MIR"
complement(11385..11572)
/rpt_family="MER69A"
complement(12325..12520)
/rpt_family="MIR"
12536..12706
/rpt_family="MIR"
12905..13529
/rpt_family="L1MA3"
13530..13842
/rpt_family="AluSp"
13843..16076
/rpt_family="L1MA3"
14517..14528
/note="<30 qual SING region"
14536..14542
/note="<30 qual SING region"
14553..14558

	/note=<30 qual SNGL region"
unure	14582. .14588
unure	/note=<30 qual SNGL region"
unure	14899. .14904
repeat_region	/note=<30 qual SNGL region"
repeat_region	16077. .16386
repeat_region	/rpt_family="AluXx"
repeat_region	16387. .16388
repeat_region	/rpt_family="LIMA3"
repeat_region	/complement(17002. .17075)
repeat_region	/rpt_family="MIR"
repeat_region	/complement(17907. .18207)
repeat_region	/rpt_family="AluSx"
repeat_region	/complement(18218. .18368)
repeat_region	/rpt_family="MIR"
repeat_region	18430. .18711
repeat_region	/rpt_family="AluSx"
repeat_region	18712. .18743
repeat_region	/rpt_family="AT_rich"
repeat_region	19184. .19685
repeat_region	/rpt_family="LTR6B"
repeat_region	/complement(19738. .19892)
repeat_region	/rpt_family="L2"
repeat_region	19894. .20155
repeat_region	/rpt_family="AluSx"
repeat_region	20167. .20189
repeat_region	/rpt_family="(CAAA)n"
repeat_region	/complement(20649. .20961)
repeat_region	/rpt_family="AluSq"
repeat_region	21297. .21411
repeat_region	/rpt_family="L2"
repeat_region	21454. .21594
repeat_region	/rpt_family="LMED"
repeat_region	21774. .22026
repeat_region	/rpt_family="MLTIE"
repeat_region	22048. .22570
repeat_region	/rpt_family="MER4A"
repeat_region	22571. .23617
repeat_region	/rpt_family="MER4A-int"
repeat_region	23618. .23651
repeat_region	/rpt_family="(TG)n"
Query Match	99.8%; Score 3730.6; DB 9; Length 191656;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 3733; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Dy	1 CTCAAAGGAGGGGTAGTCAGACGTGACGAGAACGATTGGAAGTATCATCAGCTATTAAGT 60
Db	171352 CTCAAAGGAGGGGTAGTCAGACGTGACGAGAACGATTGGAAGTATCATCAGCTATTAAGT 1712933
Dy	61 GGCAATTTAAAGCCCTCGACTAGTGAAGATTACCAAGAAGTGAAGTAGAGAGAAGA 120
Db	171292 GGCAATTTAAAGCCCTCGACTAGTGAAGATTACCAAGAAGTGAAGTAGAGAGAAGA 1712333
Dy	121 GAAGAGGCCCAAGTAGGGGATTCACAIVATTTAGATATCAGGTTGAAGAAAAGTAGTIC 180
Db	171232 GAAGAGGCCCAAGTAGGGGATTCACAIVATTTAGATATCAGGTTGAAGAAAAGTAGTIC 1711733
Dy	181 AAAAAAGTAGAGGAATCTGGGAGATCAGGTGCACAGAACCAAGTTCATAAAAAA 240
Db	171172 AAAAAAGTAGAGGAATCTGGGAGATCAGGTGCACAGAACCAAGTTCATAAAAAA 1711133
Dy	241 GACATTTAAAGAGGAAGTAGTAGAGAGAGTCAGTGCCTCTGAGAGGTAGGGTCAGAT 300
Db	171112 GACATTTAAAGAGGAAGTAGTAGAGAGAGTCAGTGCCTCTGAGAGGTAGGGTCAGAT 1710533
Dy	301 GAGAACAGAGATTGACCATGAGATTTGCCAATTGGAAGATTCAGCAACTGAGTAAG 360
Db	171052 GAGAACAGAGATTGACCATGAGATTTGCCAATTGGAAGATTCAGCAACTGAGTAAG 1709933
Dy	361 AACATTTCAATGGTTGAGCGGAACAAGCTGTAAATTGGAAGGATTTGAGGAAAAAGAC 420
Db	170992 AACATTTCAATGGTTGAGCGGAACAAGCTGTAAATTGGAAGGATTTGAGGAAAAAGAC 1709333

QY	421	AAATGGGAGCCTAGATATATCTCTTAATTAAGTTGTGTGAAAAAGAGAAAGAAAAACGGG	480
Db	170932	AAATGGGAGCCTAGATATATCTCTTAATTAAGTTGTGTGAAAAAGAGAAAGAAAAACGGG	170873
QY	481	GTGCTAGCCCAAGCTACTCTCTCACTCTTCACACCACTCATAGAGAGAGACTGGAGAACAC	540
Db	170872	GTGCTAGCCCAAGCTACTCTCTCACTCTTCACACCACTCATAGAGAGAGACTGGAGAACAC	170813
QY	541	AGCCAGAGTGAACAACTCAGTAGAAGTGTGCTTCTTTTAAGTTCTGACACTGTAT	600
Db	170812	AGCCAGAGTGAACAACTCAGTAGAAGTGTGCTTCTTTTAAGTTCTGACACTGTAT	170753
QY	601	TTCAATTATCTTAACCGGATCTCTGTAAATGGAACACTGAATTCCTTAGGAGTGTGCCCC	660
Db	170752	TTCAATTATCTTAACCGGATCTCTGTAAATGGAACACTGAATTCCTTAGGAGTGTGCCCC	170693
QY	661	CAACCCCAATGATGTTGGCTTACCTGGAACCTTAACCACTGTTTTCACACCTGCTTTC	720
Db	170692	CAACCCCAATGATGTTGGCTTACCTGGAACCTTAACCACTGTTTTCACACCTGCTTTC	170653
QY	721	TTTCAGGCACTGTGTGATTCAGATTTCAGCCAGGCAAGTGCACCAATTGCTGACAA	780
Db	170652	TTTCAGGCACTGTGTGATTCAGATTTCAGCCAGGCAAGTGCACCAATTGCTGACAA	170573
QY	781	GTCTTGCTCTATTTCTCTCTCTCACCTGCGCTCTTCATCTTGCGCTCTGTATGCACTTCT	840
Db	170572	GTCTTGCTCTATTTCTCTCTCTCACCTGCGCTCTTCATCTTGCGCTCTGTATGCACTTCT	170513
QY	841	CTCCCTCATAGCACTATTCGCACTTGATCATCACTAGCCCTCTCTGCGCTGGGCTCTGCG	900
Db	170512	CTCCCTCATAGCACTATTCGCACTTGATCATCACTAGCCCTCTCTGCGCTGGGCTCTGCG	170453
QY	901	CAGCGGCTTATGAGCAACTATGATTTCAACAGGAACTCACTAGCGGTCTGACAC	960
Db	170452	CAGCGGCTTATGAGCAACTATGATTTCAACAGGAACTCACTAGCGGTCTGACAC	170393
QY	961	AGCCGGATGCTGTGCATAGCTGTGCGCCATTTGCTCATCAGTGGCCCCAGGCTCTTCAGCA	1020
Db	170392	AGCCGGATGCTGTGCATAGCTGTGCGCCATTTGCTCATCAGTGGCCCCAGGCTCTTCAGCA	170333
QY	1021	GTGATAGCTGTCTTTTCTCTGGGAAAGAAACCTATCTGTGCTCAAGTGTATATGGGGGA	1080
Db	170332	GTGATAGCTGTCTTTTCTCTGGGAAAGAAACCTATCTGTGCTCAAGTGTATATGGGGGA	170273
QY	1081	GAGGCTTGAAGTAGAGACTGGGACAAGATATCAACTCTGATTTATTAACCATCTTTG	1140
Db	170272	GAGGCTTGAAGTAGAGACTGGGACAAGATATCAACTCTGATTTATTAACCATCTTTG	170213
QY	1141	TCCTTCAGGGCACTCAAGGATATATGTCTTCTGACAAAGGAGGCTATATACCTATGAAGCG	1200
Db	170212	TCCTTCAGGGCACTCAAGGATATATGTCTTCTGACAAAGGAGGCTATATACCTATGAAGCG	170153
QY	1201	GTATATCCGAACGGCTGGAAGAAAGATCGGGAACCCCTCATATGGGATATCTCTGGAATCTG	1260
Db	170152	GTATATCCGAACGGCTGGAAGAAAGATCGGGAACCCCTCATATGGGATATCTCTGGAATCTG	170093
QY	1261	TGATATGCGGCTTATATCTGCGCTGGGCTCTTCTCGGCTCAATATCAATGAGGATGAGGGG	1320
Db	170092	TGATATGCGGCTTATATCTGCGCTGGGCTCTTCTCGGCTCAATATCAATGAGGATGAGGGG	170033
QY	1321	TTCTGAGTCTTAGAGGGCAGCTTGTCTGTACTCTGTCTGTGGCATATAGATCCCAACAG	1380
Db	170032	TTCTGAGTCTTAGAGGGCAGCTTGTCTGTACTCTGTCTGTGGCATATAGATCCCAACAG	169973
QY	1381	GGCATAGAAAGGCTTAGGTCAAGATCCCAAGGCAATGAAAGGCTTAGGTCAAGATCCCC	1440
Db	169972	GGCATAGAAAGGCTTAGGTCAAGATCCCAAGGCAATGAAAGGCTTAGGTCAAGATCCCC	169913
QY	1441	ATGACATAGGAAGCATGCTATGTTTGTGGCTTCCCAAGGACGGCGCTGTGGTGGCT	1500
Db	169912	ATGACATAGGAAGCATGCTATGTTTGTGGCTTCCCAAGGACGGCGCTGTGGTGGCT	169853

QY	1501	3SACCTGAAGTCAGAGAGCCCAAGCCACGTCGACAGAGCTTCCTGGCCCACTGAAAGGT	1560
Db	169852	GGACCTGAAGTCAGAGAGCCCAAGCCACGTCGACAGAGCTTCCTGGCCCACTGAAAGGT	169793
QY	1561	AGACCGAGCCTTGTCGTCATGAAAAGTCCCTTGCCCTTAATCATGTCGCGCAATGGTCC	1620
Db	169792	AGACCGAGCCTTGTCGTCATGAAAAGTCCCTTGCCCTTAATCATGTCGCGCAATGGTCC	169733
QY	1621	CGGCTTGTAACCTCATTCACATGTCGCCAATTTGTAATCTGTAACGTCATGTCGAAACTGAA	1680
Db	169732	CGGCTTGTAACCTCATTCACATGTCGCCAATTTGTAATCTGTAACGTCATGTCGAAACTGAA	169673
QY	1681	TGCACGCAAGGCCCTTGCGCAACCCCAAGATGACCAAGTCCTGCGGCTGCACTACATG	1740
Db	169672	TGCACGCAAGGCCCTTGCGCAACCCCAAGATGACCAAGTCCTGCGGCTGCACTACATG	169613
QY	1741	AGGGGCTTCGACATGATGTCGCTGCGCCCACTCTAGTTCCTCATATATAAGACA	1800
Db	169612	AGGGGCTTCGACATGATGTCGCTGCGCCCACTCTAGTTCCTCATATATAAGACA	169553
QY	1801	GATTGCTTCCTGCTTCACCTGAGGGGCTTCGACATGATGTCGTCCTGCGCCCACT	1860
Db	169552	GATTGCTTCCTGCTTCACCTGAGGGGCTTCGACATGATGTCGTCCTGCGCCCACT	169493
QY	1861	CCCCGTTTCATATAATAAGACAGATTGCTTCCTTCACCTGAAATCAAGGACCTTGTCG	1920
Db	169492	CCCCGTTTCATATAATAAGACAGATTGCTTCCTTCACCTGAAATCAAGGACCTTGTCG	169433
QY	1921	TGAACAATCTCTTCCTTCCTTGAGTTGAAAAGTTAGCACTTCCTTCCTTGAGGGTCGAGC	1980
Db	169432	TGAACAATCTCTTCCTTCCTTGAGTTGAAAAGTTAGCACTTCCTTCCTTGAGGGTCGAGC	169373
QY	1981	TCAAACAAGGCTGTGAGAAAACAAGGAGGGAGCACTAAGGGCAAACTATCTTCGCG	2040
Db	169372	TCAAACAAGGCTGTGAGAAAACAAGGAGGGAGCACTAAGGGCAAACTATCTTCGCG	169313
QY	2041	AGATGATTCCTTAAGTCGTCAGATCATATAACCTAGCTCTTCCTTGACACTATCTACATAGTGG	2100
Db	169312	AGATGATTCCTTAAGTCGTCAGATCATATAACCTAGCTCTTCCTTGACACTATCTACATAGTGG	169253
QY	2101	GGGAAGAAGAACCAAGTCGGAAGAGAGACGCTGAGTTTATACAGCAAGTAAGAAGTGG	2160
Db	169252	GGGAAGAAGAACCAAGTCGGAAGAGAGACGCTGAGTTTATACAGCAAGTAAGAAGTGG	169193
QY	2161	AGCTNAGACTCTGATTCAACTTGCTGTGATGATGGCCAAACCCAGCCGAGAGGCATCAG	2220
Db	169192	AGCTNAGACTCTGATTCAACTTGCTGTGATGATGGCCAAACCCAGCCGAGAGGCATCAG	169133
QY	2221	AACACAAGGGCCTGGGGGCACTATGATGATGTCGAAAAGAGATTGGCTCAAGATGTGGGGT	2280
Db	169132	AACACAAGGGCCTGGGGGCACTATGATGATGTCGAAAAGAGATTGGCTCAAGATGTGGGGT	169073
QY	2281	AGAGAGTCCAATCTGGGGGACCTCAAAATTATGGTTCTGGGGTGAATTCAAGTAACCACTC	2340
Db	169072	AGAGAGTCCAATCTGGGGGACCTCAAAATTATGGTTCTGGGGTGAATTCAAGTAACCACTC	169013
QY	2341	ATGCGTTGTGTCATGATGATTAGGCATGACAAAGTGAATGAAGTTGAAGTGGGAAACA	2400
Db	169012	ATGCGTTGTGTCATGATGATTAGGCATGACAAAGTGAATGAAGTTGAAGTGGGAAACA	168953
QY	2401	GAATATACACACAGCTGTGTGTACAGAGGCAAGCTGAGAGAGAGAGAAAGATGAAATGGCA	2460
Db	168952	GAATATACACACAGCTGTGTGTACAGAGGCAAGCTGAGAGAGAGAGAAAGATGAAATGGCA	168893
QY	2461	CCATGAGACATTTGTCAGAACACAGTCCTGGGAGTCTTGCTGGAGGCTCAGAGACTTT	2520
Db	168892	CCATGAGACATTTGTCAGAACACAGTCCTGGGAGTCTTGCTGGAGGCTCAGAGACTTT	168833
QY	2521	GCTGGCACAAGAGATCTGGCTTACCCAAATTAGCTCTGCGGATCTGCAACATCTAGACC	2580
Db	168832	GCTGGCACAAGAGATCTGGCTTACCCAAATTAGCTCTGCGGATCTGCAACATCTAGACC	168773
QY	2581	AGCAATGTCACCTGGCAGAGAGGTGCAAGTCTGGTTATTTTCTGGTCATATAACTGGTG	2640

Db	168772	AGCAATATCTACTGGCAGAGAGTTGCAGTGTGTTATTTTCTGTGCTATTAACCTGGTG	168713
OY	2641	AAGCTTTGGGTTCCAAATTTTGCTGACAGCTGTTTTAACTGGAAATTGGCCTTAGACTATA	2700
Db	168712	AAGGCTTTGGGTTCCAAATTTTGCTGACAGCTGTTTTAACTGGAAATTGGCCTTAGACTATA	168653
OY	2701	GGTAGCTATGTCTCAGACAGAGCCCTATTTCCGCCATCCCTTTAACAACCAGCTAGAGTT	2760
Db	168652	GGTAGCTATGTCTCAGACAGAGCCCTGTGTCTCCACTGTCTTACCAACCAGCTAGAGTT	168593
OY	2761	GGAGGCTGGCTGTTTACAGCTCCAAAAAATAGCTGAGTTTCCAGACAGAGGCCCTTATT	2820
Db	168552	GGAGGCTGGCTGTTTACAGCTCCAAAAAATAGCTGAGTTTCCAGACAGAGGCCCTTATT	168533
OY	2821	CTGAGCTTCCGTTGCTTAGCCTCATTTTTCTTTCTGTGTAATTAAGACAATGCCACCCA	2880
Db	168512	CTGAGCTTCCGTTGCTTAGCCTCATTTTTCTTTCTGTGTAATTAAGACAATGCCACCCA	168473
OY	2881	CCTTCCAGTGAACATGAATATTAAGACTCAAACCATCCCTTGAACGTCTTGGGAGGGGC	2940
Db	168472	CCTTCCAGTGAACATGAATATTAAGACTCAAACCATCCCTTGAACGTCTTGGGAGGGGC	168413
OY	2941	TCTGGAGCTAAGCCAGACTGAGGCTCAGGCTCATGATGATCTGAGAGTCAAGCCCTCC	3000
Db	168412	TCTGGAGCTAAGCCAGACTGAGGCTCATGATGATCTGAGAGTCAAGCCCTCC	168353
OY	3001	AACCTGTACAGCAATTTGCTCCGTAGAGACTTTGATGGGTAGATAGCTTAAACAAGCTCT	3060
Db	168352	AACCTGTACAGCAATTTGCTCCGTAGAGACTTTGATGGGTAGATAGCTTAAACAAGCTCT	168293
OY	3061	GACTGTCAACAAGGCTTTGTATCTGGAGGCCAGGCTTATAGATGGCTCCAGCTTAAAG	3120
Db	168292	GACTGTCAACAAGGCTTTGTATCTGGAGGCCAGGCTTATAGATGGCTCCAGCTTAAAG	168233
OY	3121	GCTGGAGACTGGGGGACAGTGTCTCAGATTAGGCTTAACTAAGGAAGTTGACTGGAGCTG	3180
Db	168232	GCTGGAGACTGGGGGACAGTGTCTCAGATTAGGCTTAACTAAGGAAGTTGACTGGAGCTG	168173
OY	3181	AGAAACAAGGTTAAGGGGCAAGACAGAGGTTGTGGTCTACTCCTTAGAGACACTTGA	3240
Db	168172	AGAAACAAGGTTAAGGGGCAAGACAGAGGTTGTGGTCTACTCCTTAGAGACACTTGA	168113
OY	3241	GCTTTACTTTTCATTCTTAATGTGTCTTGGATGGCTACCTCAAGGGTGGCTGTAG	3300
Db	168112	GCTTTACTTTTCATTCTTAATGTGTCTTGGATGGCTACCTCAAGGGTGGCTGTAG	168053
OY	3301	TCTTAAGGGGTGAGACAAGGACAAGAGTTACAGTCTGATCGCTTATCAAGTTATGACATA	3360
Db	168052	TCTTAAGGGGTGAGACAAGGACAAGAGTTACAGTCTGATCGCTTATCAAGTTATGACATA	167993
OY	3361	CACCTGGACCACTGTCTCATATGCGCAGGAGCCTTAGAGGTCTTAAACAGTTATCCA	3420
Db	167992	CACCTGGACCACTGTCTCATATGCGCAGGAGCCTTAGAGGTCTTAAACAGTTATCCA	167933
OY	3421	CAACTGTGATACCAAGGTTAACTTTCTCTGTTTTCAAGGCAAGGAGTACTAAGTCTC	3480
Db	167932	CAACTGTGATACCAAGGTTAACTTTCTCTGTTTTCAAGGCAAGGAGTACTAAGTCTC	167873
OY	3481	CCCTTTCTCCCTTCTCTCCACGTGTCTCTTGCAGAGGAATCCTTAGCTGTGCTCAAGGG	3540
Db	167872	CCCTTTCTCCCTTCTCTCCACGTGTCTCTTGCAGAGGAATCCTTAGCTGTGCTCAAGGG	167813
OY	3541	AACCTCCAGAATGGTTGTTCAGTCAAGTTAAGGCTCTATTAAGAAATATCTTAGAGT	3600
Db	167812	AACCTCCAGAATGGTTGTTCAGTCAAGTTAAGGCTCTATTAAGAAATATCTTAGAGT	167753
OY	3601	GGGTATATCTATCAGCAATAGGAATTTATTTGTTCAATTTCTGAGAGCTGGAATTCAG	3660
Db	167752	GGGTATATCTATCAGCAATAGGAATTTATTTGTTCAATTTCTGAGAGCTGGAATTCAG	167693
OY	3661	ATCAAGGCTCAGACAGTTCAAGTGTCTGTGAGTGTCTTCTGTTGGAAATGGCACCC	3720

Db 167692 ATCAGAGCTCCAGAGTCTGCTGAGNGCTTCTTCGTCGACAGATGACACC 167633

QY 3721 TTTTGTGTTGTTCTCA 3737

Db 167632 TTTTGTGTTGTTCTCA 167616

RESULT 2

AC084337/c 112295 bp DNA linear PRI 01-JUN-2002

LOCUS Homo sapiens chromosome 11, clone CTD-2010116, complete sequence.

DEFINITION AC084337

AC084337.7 GI:21306722

VERSION HTG.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 112295)

Birken, B., Linton, L., Nusbaum, C., and Lander, E.

Homo sapiens chromosome 11, clone CTD-2010116

Unpublished

2 (bases 1 to 112295)

Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Bede, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlino, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Macdonald, P., Marguis, N., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Morow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivar, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (25-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 112295)

Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Cooke, P., Dearlino, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goid, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Ton, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rieck, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 112295)

AUTHORS

Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Cooke, P., Dearlino, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goid, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Ton, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rieck, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 1, 2002 this sequence version replaced gi:19683126.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIMR

Web site: <http://www-seg.wi.mit.edu>

Contact: sequence submissions@genome.wi.mit.edu

----- Project Information

Center project name: L10956

Center clone name: 2010_I_16

Only the last 112.3 kb of this clone are being submitted.

The remainder overlaps AC068733 [WIGR project L10266].

Location/Qualifiers

1. 112295

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="CTD-2010116"

/clone_11b="CTD1 Human BAC"

complement(949..1264)

/rpt_family="AluSg"

repeat_region

1269..1568

/rpt_family="L2"

complement(2334..2584)

repeat_region

/rpt_family="AluSg"

complement(2429..2433)

/note="<30 qual SINGL region"

unsure

complement(2462..2466)

/note="<30 qual SINGL region"

unsure

complement(2478..2503)

/note="<30 qual SINGL region"

unsure

complement(2508..2537)

/note="<30 qual SINGL region"

unsure

complement(2577..2583)

/note="<30 qual SINGL region"

repeat_region

2772..3185

/rpt_family="L2"

complement(4149..4288)

repeat_region

/rpt_family="L2"

4268..4443

/rpt_family="L2"

repeat_region

4453..4524

```

repeat_region      /rpt_family="L2"
                    4563. .5531
                    /rpt_family="L2"
repeat_region      complement(5560. .5857)
                    /rpt_family="ALUSX"
repeat_region      5861. .6355
                    /rpt_family="L2"
repeat_region      complement(6378. .7246)
                    /rpt_family="L2"
repeat_region      7898. .7942
                    /rpt_family="GC_rich"
repeat_region      8243. .8254
                    /rpt_family="GC_rich"
repeat_region      complement(8355. .8589)
                    /rpt_family="MIR"
repeat_region      complement(8875. .9148)
                    /rpt_family="ALUto"
repeat_region      9264. .9349
                    /rpt_family="L2"
repeat_region      complement(9361. .9480)
                    /rpt_family="LMC5"
repeat_region      complement(9584. .9639)
                    /rpt_family="L2"
repeat_region      9656. .9676
                    /rpt_family="(TTCA)n"
repeat_region      complement(9677. .10005)
                    /rpt_family="L2"
repeat_region      10006. .10186
                    /rpt_family="MSTA"
repeat_region      10187. .10496
                    /rpt_family="ALUSq"
repeat_region      10497. .10671
                    /rpt_family="MSTA"
repeat_region      complement(10672. .10981)
                    /rpt_family="ALUY"
repeat_region      10982. .11052
                    /rpt_family="MSTA"
repeat_region      complement(11053. .11484)
                    /rpt_family="L2"
repeat_region      complement(12337. .12474)
                    /rpt_family="PLAM C"
repeat_region      complement(12507. .12954)
                    /rpt_family="L2"
repeat_region      12998. .13714
                    /rpt_family="L1MA7"
repeat_region      complement(14073. .14372)
                    /rpt_family="ALUSg"
repeat_region      14466. .14974
                    /rpt_family="MLT1K"
repeat_region      complement(14975. .15266)
                    /rpt_family="ALUSX"
repeat_region      15267. .15301
                    /rpt_family="MLT1K"
repeat_region      15337. .15440
                    /rpt_family="LTR50"
repeat_region      15632. .15817
                    /rpt_family="LTR33A"
repeat_region      complement(15824. .15945)
                    /rpt_family="MIR"
repeat_region      16067. .16090
                    /rpt_family="AT_rich"
repeat_region      complement(16091. .16371)
                    /rpt_family="ALUto"

Query Match      99.8%; Score 3729; DB 9; length 112295;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3732; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Db      21705 GGCATTAAAGCCCTGAGCTAGGTGAGATTACCAAGAAAGTAAAGAGAGAAGA 21646
Qy      121 GAAGAGGCCCAAGTATGGGGATTCCAATATTATGATATCAGTTGAAGAAAGGTGTC 180
Db      21645 GAAGAGGCCCAAGTATGGGGATTCCAATATTATGATATCAGTTGAAGAAAGGTGTC 21586
Qy      181 AAAAAATTAAGAGATATCTGGAGAGTCAAGAGTTCACAAAGTCCAAAAA 240
Db      21585 AAAAAATTAAGAGATATCTGGAGAGTCAAGAGTTCACAAAGTCCAAAAA 21526
Qy      241 GACATTTAAAGAGAGAAAGTATGAGAGTCAAGTCTCTGAGAGGTGATGAT 300
Db      21525 GACATTTAAAGAGAGAAAGTATGAGAGTCAAGTCTCTGAGAGGTGATGAT 21466
Qy      301 GAGAACAGAAATGACATGAGATTTCGCAATTGGAATTACTAGCACTGATPAG 360
Db      21465 GAGAACAGAAATGACATGAGATTTCGCAATTGGAATTACTAGCACTGATPAG 21406
Qy      361 AACATTTCAATGTTGAGGAGAAACAGAAAGTAAATTAAGAGATTGAGAAAAAGAC 420
Db      21405 AACATTTCAATGTTGAGGAGAAACAGAAAGTAAATTAAGAGATTGAGAAAAAGAC 21346
Qy      421 AAATGGAGCTTGATATCTTAAATAGTTGTGAAAAAGAGAAAAACGGG 480
Db      21345 AAATGGAGCTTGATATCTTAAATAGTTGTGAAAAAGAGAAAAACGGG 21286
Qy      481 GTGCTAGCCAGATCTCCCTCACTTCCACACACCTGAGGAGAGATGAGAACAC 540
Db      21285 GTGCTAGCCAGATCTCCCTCACTTCCACACACCTGAGGAGAGATGAGAACAC 21226
Qy      541 AGCCAGAGTGAAGAACTTCAAGTAAAGTGTCTCTTTTAAGTCTGACACTGTAT 600
Db      21225 AGCCAGAGTGAAGAACTTCAAGTAAAGTGTCTCTTTTAAGTCTGACACTGTAT 21166
Qy      601 TTGATTAATCTAATACCGCATCTCTGACATGACACCTGAAATCTTGAAGTCCCGC 660
Db      21165 TTGATTAATCTAATACCGCATCTCTGACATGACACCTGAAATCTTGAAGTCCCGC 21106
Qy      661 CAAGCCAGATGTTGGCCTTACCTGAGAACTTGAAGCACTTTTCAACCTTGCTTTC 720
Db      21105 CAAGCCAGATGTTGGCCTTACCTGAGAACTTGAAGCACTTTTCAACCTTGCTTTC 21046
Qy      721 TTTCAAGGACCTGCTGATTCAGTTTCAGGAGGAGCAAGTGTCCCAATGTCTGACCA 780
Db      21045 TTTCAAGGACCTGCTGATTCAGTTTCAGGAGGAGCAAGTGTCCCAATGTCTGACCA 20986
Qy      781 GTCTTGCTAATTTCTCTTCTCACTGAGCTCTTGCATTTGGCTTGGATGCAATCT 840
Db      20985 GTCTTGCTAATTTCTCTTCTCACTGAGCTCTTGCATTTGGCTTGGATGCAATCT 20926
Qy      841 CTCCTCTCATGACTAATTTCTGATTCATCACTAGAGCTCTTCTGAGCTGAGCTTGC 900
Db      20925 CTCCTCTCATGACTAATTTCTGATTCATCACTAGAGCTCTTCTGAGCTGAGCTTGC 20866
Qy      901 CAGCGGCCCTAGAGCAACTATGATATTCACAGGAGCCACCTACTGCGCTTGACACC 960
Db      20865 CAGCGGCCCTAGAGCAACTATGATATTCACAGGAGCCACCTACTGCGCTTGACACC 20806
Qy      961 AGCCGGAGTGTGCTGAGTACTGAGCCATTTGCTCATCATGAGGCTCCCTTACGCA 1020
Db      20805 AGCCGGAGTGTGCTGAGTACTGAGCCATTTGCTCATCATGAGGCTCCCTTACGCA 20746
Qy      1021 GTGATGCTGCTTTCTCTGGAGAAAGAACTATCTGCTCAGAGGTGATATGGGGGA 1080
Db      20745 GTGATGCTGCTTTCTCTGGAGAAAGAACTATCTGCTCAGAGGTGATATGGGGGA 20686
Qy      1081 GAGCTTGAAGTGAAGTGGAGCAAGCATATCCATCTGTATTTATTAACATCTTTTG 1140
Db      20685 GAGCTTGAAGTGAAGTGGAGCAAGCATATCCATCTGTATTTATTAACATCTTTTG 20626
Qy      1141 TCCTCAGGCGACCCAGTATATGTCTTCTGCAAAAGGAGGCTATACCTATGTAAGCG 1200

```

Dp	20625	TCCCTCAGGGCACCAGGTAATATGTCTTCTGACAAAGGAGCTATACCCTATGTAAGC	20586
Qy	1201	GTTATCCGAAGCGGCTGAGAAAGAAAGTCGGAGCCCTCATAGGATTAATCTTGACTGTG	1260
Dp	20565	GTTATCCGAAGCGGCTGAGAAAGAAAGTCGGAGCCCTCATAGGATTAATCTTGACTGTG	20506
Qy	1261	TGAGATGGGCGCTTTATTTGCGCCCTGGGTCCTTCTGGCTCCATATCATGGCAGGTAGGGGC	1320
Dp	20505	TGAGATGGGCGCTTTATTTGCGCCCTGGGTCCTTCTGGCTCCATATCATGGCAGGTAGGGGC	20446
Qy	1321	TTTCGGGCTGTCTGAGGGGCAAGCTGTTCTGCTACCTGTCGTGGGATATATCCCAACAG	1380
Dp	20445	TTTCGGGCTGTCTGAGGGGCAAGCTGTTCTGCTACCTGTCGTGGGATATATCCCAACAG	20386
Qy	1381	GGCATGGAAGGCGCTAGTCAAGATCCCAAGGCAATGAGAAGGCTAGGTCAAGATCCCC	1440
Dp	20385	GGCATGGAAGGCGCTAGTCAAGATCCCAAGGCAATGAGAAGGCTAGGTCAAGATCCCC	20326
Qy	1441	ATGACATGGAAGCCATGCTATGTTTGGTGCCTTCTCCAGAGCGGCGGCTGTGGTGCCT	1500
Dp	20325	ATGACATGGAAGCCATGCTATGTTTGGTGCCTTCTCCAGAGCGGCGGCTGTGGTGCCT	20266
Qy	1501	GGACCTGGAATCAAGAGGCCCAAGCAACGTGGACAGAGGCTTCCCTGGCCCCATGAGAAGGT	1560
Dp	20265	GGACCTGGAATCAAGAGGCCCAAGCAACGTGGACAGAGGCTTCCCTGGCCCCATGAGAAGGT	20206
Qy	1561	AGACGAGAGCTTGTTGTATGGAAGAAAGTCCCTTGGGCCCTTAATCAATGTTCCGCAATAGTCTC	1620
Dp	20205	AGACGAGAGCTTGTTGTATGGAAGAAAGTCCCTTGGGCCCTTAATCAATGTTCCGCAATAGTCTC	20146
Qy	1621	CGGCTTGTAACCTCAATCAGTGGTCCCAATTTGTACTGCTACAGTGAATGAGAACTGAA	1680
Dp	20145	CGGCTTGTAACCTCAATCAGTGGTCCCAATTTGTACTGCTACAGTGAATGAGAACTGAA	20086
Qy	1681	TGCAGCCCAAGGCCCTTCCGGAACCCCAAGATGTGACCAAGTCTCTGGGCTGCATCACTG	1740
Dp	20085	TGCAGCCCAAGGCCCTTCCGGAACCCCAAGATGTGACCAAGTCTCTGGGCTGCATCACTG	20026
Qy	1741	AGGGGCGCTTCTGACATGAGTCTGGCGCGGCCCACTCCTAGTTCCTCATTAATAAAGACA	1800
Dp	20025	AGGGGCGCTTCTGACATGAGTCTGGCGCGGCCCACTCCTAGTTCCTCATTAATAAAGACA	19986
Qy	1801	GATTGCTTCTTGGCTTCTCACTGAGGGGCGCTTCTGACATGAGTCTGGCGGCCCACT	1860
Dp	19985	GATTGCTTCTTGGCTTCTCACTGAGGGGCGCTTCTGACATGAGTCTGGCGGCCCACT	19906
Qy	1861	CCCCAGTTTCTCATTAATAAAGACAGATTGCTTCTTCACTTGAATCAAGGACCTTGGTGTG	1920
Dp	19905	CCCCAGTTTCTCATTAATAAAGACAGATTGCTTCTTCACTTGAATCAAGGACCTTGGTGTG	19846
Qy	1921	TGAAACAATCTTCTTCTTGAAGTGAAGATTAGCACTTCTCTTGAAGGGTGTGAGC	1980
Dp	19845	TGAAACAATCTTCTTCTTGAAGTGAAGATTAGCACTTCTCTTGAAGGGTGTGAGC	19786
Qy	1981	TCAAACAAGCGTGTGAGAAACAAGGAGGGGAGCACTAAGGGGCAAACTATCTCTGCGC	2040
Dp	19785	TCAAACAAGCGTGTGAGAAACAAGGAGGGGAGCACTAAGGGGCAAACTATCTCTGCGC	19726
Qy	2041	AGATGATTTCTTAGGTCCAGATCATTAACTAAGTCTTGTGACACTATCTACATAGTGGG	2100
Dp	19725	AGATGATTTCTTAGGTCCAGATCATTAACTAAGTCTTGTGACACTATCTACATAGTGGG	19666
Qy	2101	GGGAAAGAGAACCAAGATCGGAAAGAGAAACAGCTGAGTTTATACACAAAGTAAGAGGTGG	2160
Dp	19665	GGGAAAGAGAACCAAGATCGGAAAGAGAAACAGCTGAGTTTATACACAAAGTAAGAGGTGG	19606
Qy	2161	AGCTAGGACTCTGATTCAACTTGTGTGATAGTGGCAACAACCAAGCGCAAGGCAATCAGA	2220
Dp	19605	AGCTAGGACTCTGATTCAACTTGTGTGATAGTGGCAACAACCAAGCGCAAGGCAATCAGA	19546
Qy	2221	AACAAACGGGCTTGGGGCAACTATGACATGTGCAAAAGAGATTGGCTCAAGATTGTGGGCT	2280
Dp	19545	AACAAACGGGCTTGGGGCAACTATGACATGTGCAAAAGAGATTGGCTCAAGATTGTGGGCT	19486

QY	2281	AGAGGCTCCAAATCTGGGGGACCTCAAAATTAAGTTCTTGGGTGATTCAAATTAACAACAATC	2340
Db	19485	AGAGGGTCCAAATCTGGGGGACCTCAAAATTAAGTTCTTGGGTGATTCAAATTAACAACAATC	19428
QY	2341	ATGACTGTGTTCACATGATTAGGCATGACAAGTGGAAATGAAGTTGAAGTGGGAAACA	2400
Db	19425	ATGACTGTGTGTGCANTGATTAAGGCATGACAAGTGGAAATGAAGTTGAAGTGGGAAACA	19366
QY	2401	GAAATTAACAACGCTGTGTGTCAAGGCGAAGCTGGAGAGAGAGAAAGAAATGAATGCA	2460
Db	19365	GAAATTAACAACGCTGTGTGTCAAGGCGAAGCTGGAGAGAGAGAAAGAAATGAATGCA	19306
QY	2461	CCATGGAGGCAATTTCGAGAAACAGTCCCTGGGAGCTTCCTGGAGGCTCAGAGACTTT	2520
Db	19305	CCATGGAGGCAATTTCGAGAAACAGTCCCTGGGAGCTTCCTGGAGGCTCAGAGACTTT	19244
QY	2521	GCTGGCAACAGAGATCTGGCTTACCCAAATTAGCTCTCGGGATCTGCAACATTAAGCC	2580
Db	19245	GCTGGCAACAGAGATCTGGCTTACCCAAATTAGCTCTCGGGATCTGCAACATCTAGACC	19186
QY	2581	AGCAAAATGTAATGGGACAGAGAGGTTGCAAGCTGTGGTTATTTCTGTCATATAAATGCGTG	2640
Db	19185	AGCAAAATGTAATGGGACAGAGAGGTTGCAAGCTGTGGTTATTTCTGTCATATAAATGCGTG	19128
QY	2641	AAGCTTTGGGTTCCAAATTTGCTGACAGCTGTTTAATCTGGAAATTGGGCTTAGACTATA	2700
Db	19125	AAGCTTTGGGTTCCAAATTTGCTGACAGCTGTTTAATCTGGAAATTGGGCTTAGACTATA	19066
QY	2701	GGTAGCTATGTCTTCAGACAAAGGCGCTTATCTCCACTGCTTTTCAACCAAGCTGAGTT	2760
Db	19065	GGTAGCTATGTCTTCAGACAAAGGCGCGTGTCTCCACTGCTTTTCAACCAAGCTGAGTT	19008
QY	2761	GGAGGCTGGCTGTGTTTCAGGCTCAAAAATAAGCTAGATTTCCAGAGAGGGGCGCTTAT	2820
Db	19005	GGAGGCTGGCTGTGTTTCAGGCTCAAAAATAAGCTAGATTTCCAGAGAGGGGCGCTTAT	18946
QY	2821	CTGAGCTTCGGTGTCTTAGGCTCATTTTCTTTTCTGTAATAATGACACAAATGCCACCA	2880
Db	18945	CTGAGCTTCGGTGTCTTAGGCTCATTTTCTTTTCTGTAATAATGACAAATGCCACCA	18886
QY	2881	CCTTCCAGTGAACAATGAATTAACATCTCAACCAATCCCTTGAACTGTCTTGGGAAGGGGC	2940
Db	18885	CCTTCCAGTGAACAATGAATTAACATCTCAACCAATCCCTTGAACTGTCTTGGGAAGGGGC	18828
QY	2941	TCTGAGATGAGACCCAGACTGTGGCTCATAGGCTCATGTGATCTGAGAGTCAAGCCCTGCC	3000
Db	18825	TCTGAGATGAGACCCAGACTGTGGCTCATAGGCTCATGTGATCTGAGAGTCAAGCCCTGCC	18766
QY	3001	AACCTGTACGCAATTGCTCCGTAGGACTTTGATGGGTAGAGTAGTACTTAACAAGCTCT	3060
Db	18765	AACCTGTACGCAATTGCTCCGTAGGACTTTGATGGGTAGAGTAGTACTTAACAAGCTCT	18708
QY	3061	GACTGTACAACAAGGCTTTGTAATGGGAGGCGAGGCTATAGATGGGCTCAGCTTAAAG	3120
Db	18705	GACTGTACAACAAGGCTTTGTAATGGGAGGCGAGGCTATAGATGGGCTCAGCTTAAAG	18644
QY	3121	GCTGGAGCTGTGGGAGCAGTGTCTCAGATTAGGGTCTAACTAGGAAGTTGACTGAGCTG	3180
Db	18645	GCTGGAGCTGTGGGAGCAGTGTCTCAGATTAGGGTCTAACTAGGAAGTTGACTGAGCTG	18588
QY	3181	AGAACAGAGATTAGGGGCCAAGACAGAGGTTGTGGGTCTACCTCTTAGAGACACTTGA	3240
Db	18585	AGAACAGAGATTAGGGGCCAAGACAGAGGTTGTGGGTCTACCTCTTAGAGACACTTGA	18528
QY	3241	GCTTTACTTTTCAATCTTAATGCTGTCTTGGATGAGCTACCCCTCAACGAGGTTGGCTGAG	3300
Db	18525	GCTTTACTTTTCAATCTTAATGCTGTCTTGGATGAGCTACCCCTCAACGAGGTTGGCTGAG	18466
QY	3301	TCTAAGGGGTGGAGACAAGACAGAGTTTCAAGGTCTGATCTTATCAAGTTCAATGACATA	3360
Db	18465	TCTAAGGGGTGGAGACAAGACAGAGTTTCAAGGTCTGATCTTATCAAGTTCAATGACATA	18406

Db	4632	TTTTTTTATCTGGAGATAAATTAAATTATATATTAAGAATGATTTTAAAGTTATTTATTTT	4691
Qy	1692	TGACTTATGATGTCCTGGACCCATTGCCAGTGATGAGATCAACGTGAGCCTCAGCAGT	1751
Db	4692	TGATTTATGATGATTTGGATTTTATTTATGTTAGTGTGAGAGATTATATGTTGACCGTTAGTAGT	4751
Qy	1752	TAGCCCACTTCTGCTTGAANAATTGCTGAAGGGGGTTGGGGGGCACGCTCCGGGAAAAG	1811
Db	4752	TAGTTTATGTAATGTTTGTGAATAATGTTTAAAGGGGGTTGGGGGGTATGTTGTGGGAAAAAG	4811
Qy	1812	GAGCTTGATGATTCAGATTTCTGTGCCAGACCCTGCAGCCTATTGTCAGTGATGAATCAGCC	1871
Db	4812	GAGTTTGGATTTTATGATTTTGTGTTAGATTTTGAATTTATTTGTAATGTAATGTAATGATTT	4871
Qy	1872	AATATGCGCTTAGTCTCTGGAGAACAGACATTTCCAGTAGAGTTGAGAGTGGGGTGGTG	1931
Db	4872	AATATTTGGTTAGTTTGTGGAGATAGTATATTTTATAGTAGAGTTGAGAGTGGGGTGGTG	4931
Qy	1932	CTGCTGCCACTCTATATATAGGAGATTCACATGCTACCCAGACGCTCTCTGCGCTCTG	1991
Db	4932	TGTGTTTAAATTTTATATATAGGAGATTTAATGTTAATGTTAATGTTTGTGTTTGTG	4991
Qy	1992	CAGCTCAGCATGCTAGGGTATCTGGAGACACCCGTTGCACTGGGGTTGGAGCCTATGC	2051
Db	4992	TAGTTTATGATGTTTATGAGATTTGGAGATATTCGTTATTTGGGGTTGTGAGATTATGT	5051
Qy	2052	TGCTCTCGCCATTGCCACCCCTCTCTCCGTAGATGAAGCTGGGACTGAAGCGAG	2111
Db	5052	TGGTTTTGTTATTTATTTTATTTTCTTCTCGTAGATGAAGTTGGATTGAAGCGAG	5111
Qy	2112	GATTGATGTTCTGGGCTAGGGTATAGGTAGGGCCAGTTTATAGGCTCGTCAAATTTGGGG	2171
Db	5112	GATTGATGTTTGGGTTAGGGTATAGGGTATAGGTTTATAGGTTTCGGTTAAATTTGGGG	5171
Qy	2172	TCAGGGGCTATGGGAAAGAGATCGGCTCCAGATGATCAAGATATATTTGTTCTCCCT	2231
Db	5172	TTAGGGGTTATGGAAAGAGATCGGTTTAAATGATTAAGATTAATTTATTTGTTTTTTT	5231
Qy	2232	AGG 2234	
Db	5232	AGG 5234	
<hr/>			
RESULT 6	AX345847/C	5234 bp	DNA linear PAT 01-FEB-2002
LOCUS	AX345847		
DEFINITION	Sequence 918 from Patent WO0200928.		
ACCESSION	AX345847		
VERSION	AX345847.1 GI:18493733		
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	artificial sequences.		
AUTHORS	1 Olek,A., Piepenbrock,C. and Berlin,K.		
TITLE	Diagnosis of diseases associated with the immune system		
JOURNAL	Patient: WO 0200928-A 918 03-JAN-2002;		
FEATURES	Epigenomics AG (DE)		
Source	location/Qualifiers		
	1..5234		
	/organism="synthetic construct"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32630"		
	/note="chemically treated genomic DNA (Homo sapiens) "		
BASE COUNT	1303 a 40 c 1091 g 2800 t		
ORIGIN			
Query Match	12.8% Score 1281.6; DB 6; Length 5234;		
Best Local Similarity	73.4% Pred. No. 0;		
Matches 1638; Conservative	0; Mismatches 594; Indels 0; Gaps 0;		
1 TCCCTCTCCCAGGACGCCACGAAATCTGTAGATTCAAGACGGGTTCTGACAGCTG	60		

Db	2234	TCCTCTCCCAACCAACCCACAAATCTATAAATTCAAACAAATTTCAACACTA	2175
Oy	61	AAGACAAGTGTGAGGAAATTCCTGATGAGATCATGGGCTCTCAGGAGGAATA	120
Db	2174	AAACAATATTATTAATAAAATTCCTAATTAATAAAATCATAAATCTCAAAAAAATAAATA	2115
Oy	121	TAAGTTTCAGAGGCTGAGAGGAAAGAAAAAGTGAAGGGAGCTTAAATTAAGTCTC	180
Db	2114	TAAATTTCAAAAACTAAAAAATAAAAAAATAAAAAAATCTTAAATTAATCTC	2055
Oy	181	CCATTGCCAACACCCAGAAAGAAGATGCGCCCTGCAGTGGGAGAAAGTGAGATGAGA	240
Db	2054	CCATTACCCAGACCCCAAAAAAATAAAACATACCCTCAATTAATAAAAAATTAATTAATAA	1995
Oy	241	CATTGGCTGAGCAGCATGGCATTTGCCAGGCTGCCAAGACTCAGAGACTCGACCTT	300
Db	1994	CATTAACTATTAACAACGATTAACATTACCAAACTACCAAAAATCTCAAAAAATTCACACTT	1935
Oy	301	GCCCACTGACTTATGAGAGGAGATGATGTTCAACACATTTTCATTCGTAAGTCAGA	360
Db	1934	ACCCACTTAACCTAATMAAAAAAATAATTAATTCACAAACATTTTCATTCGTAATTCAAA	1875
Oy	361	GAGGACATATGAGCGCTGATGGCAGAGGCGCTGGTGCATGTTGTCAGAGGTTCCGGATG	420
Db	1874	AAAAACATTAACCTAATTAACAAAAACCTAATTAACATTAATTCACAAAAATTCGAAATA	1815
Oy	421	TGTGTTTTCTGTGTAAGAAAGAACTTCGACAGATGAAAAAGGATCTGAGACTTTGGTA	480
Db	1814	TATATTTTCCATTAATAAAAAAATTTGGCAAAATTAATAAAAAAATCTTAACCTTTATA	1755
Oy	481	AGATTATATATGGAAGTGTCAAGGGGTTCGAGCCATCTGAGAGGATCAGGCGCTTTCA	540
Db	1754	AAATTAATATATAAACATCAAAATCTTAAACCATCTATMAAAATCAAAACCTTTCA	1695
Oy	541	GCTTGGCTAGGAGCAGGGGTCCTGGAACCTTATCCTGCGCCATAGCTGAGCTGCCCA	600
Db	1694	ACCTTAACCTMAAAACAAAAATCCTAAAACTTCAACCTTAACCAATTAATCTAACCA	1635
Oy	601	TAATCTTTTTCTGACTCACTAGCAAAATCTCAACAGAAATGGGGCAGCTTTGGGAGTGG	660
Db	1634	TAAATCTTTTCTAATCTCACTTAACAAATCTCACACAAAAATMAAACACTTTAAATAA	1575
Oy	661	GCCCAAGAAAGTCTGAGGATGACAGGTGAGATCCACAGAAAGATGATGTGGGCGCAG	720
Db	1574	ACCCAAAAAATACTAAATAATTAACAATAATTAATTAATTAATTAATTAATTAATTAATTA	1515
Oy	721	ACACTGAGAGAGAAACAGACTGTCAATTAAGAGGGGTCTGTGACATCTTAAGATCTCAAT	780
Db	1514	ACACTTAATAAAAAAACAACAACTTCAAAATTAATAAAAGTCTATATACCTTAATCTCAAT	1455
Oy	781	ATGCTCTACTACATAACCTACCCCAATTCCTAATATTTCTCTACCTTAGAGGGGGGAA	840
Db	1454	ATACTCTACTACATAACCTACCCCAATTCCTAATATTTCTCTACCTTAATAAAAAAATAA	1395
Oy	841	ATTGTCAAAATTTGGCTCAACCTAGCAACACTACTCACTAGTACTTGAATGCAATTTTG	900
Db	1394	ATTATCAAAATTTAATCTACAAACCTAACAACACTACTCAATTAATTAATAATTAATTTTA	1335
Oy	901	CATTTTTCATTTCAACAATATTTCTGGAACACTCTTATATGCGCAGACTATTTTGA	960
Db	1334	CATTTTTTCACTCAACAAATTTCTTAAACAACCTTATATATCCTTAATACCTATTTTAA	1275
Oy	961	GAGTCAGGAGATATTAATGTTAAACAAGCAGGAAACAAAGCAAAAGCAACAACA	1020
Db	1274	AAATCAAAATATATTAATTAATTAACAACAACAACAACAACAACAACAACAACAACA	1215
Oy	1021	TCACCAATATAGTACAGATGAAGAATTTCAAGTTTATAGTAATTAATAAACAAGC	1080
Db	1214	TCACCAATATATTAACAACAATTAATAAAATTTCAAAATTTTATATAATTAATAAACAAC	1155
Oy	1081	AAGGCTGAAATGGCTAGATAGCGGCTCAGAAAGGCTTCATTGCAAGTACGATTT	1140

```

Db 1154 AAAATCTAAATACTAAATAATAATCAAAAAAATCTGATTAATAATAATCAATT 1095
QY 1141 AACGAGAGTCAGTGAATAATATGTGAATTCAGTTAGATTTCTATTTGTTCTGGGT 1200
Db 1094 AAACAAAATCAACATAAATATTTATTAATTCATTAATTCATTAATTTCTAAAT 1035
QY 1201 GGTAAATAAAGCTTTTCCCAAGGTGAAAATACCAAGAAAAGATTAATTAAGTAG 1260
Db 1034 AATTAATAAATCACTTTTCCCCCAATAAATAAATACTACAAAATAAATAATTAATA 975
QY 1261 TGGTGGTCTCTGAGAGAGACACCTCTGTTTGTGCTCATTAATCTGACCTTC 1320
Db 974 TAATAATACCTCTAATAAATAAATAAATCTTAATTTCTACCTCATTAATCAACCTTC 915
QY 1321 ACTTCAGGAGCTTTTTCGAAGCCCTTTCCAGTCAGGAGGAGGAGGCTGGGATG 1380
Db 914 ACTTCGAACACTTTTTCGAAGCCCTTTACCATCAAAAAAATCAAAAAATCAAACTA 855
QY 1381 GGGCTTGAACATTTGACACAGTGAACATTAATTTCCCAAGTCACTAGCCCAAGGT 1440
Db 854 AAATTAATCAATTAACACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 795
QY 1441 AAAGCTGAAGAGCTGGGATGCCCCGAGAAAGCCCTGATGAGCTTGAAGAAAGCTG 1500
Db 794 AAACCTAATAAATACTTAACATACCCCAAAAAAATCTTAATAAATCAAAAACTA 735
QY 1501 TTCTGAGATTTCTAGTAAGTTATCTGTGTGTGTGTGTTAATAAGTAAGTAT 1560
Db 734 TTCTGTAATATTTCTTAATTAATTTCTATATTAATTAATTAATTAATTAATTAAT 675
QY 1561 TGTGTCTAGTGTGCTTAGAGCAGGCTTGAACAGTAACAGCAATTAATTTCCCT 1620
Db 674 TACTATCTTAACCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 615
QY 1621 CCTTTCTCACTCCCAATTTGAGATTAATCAATCAAAAAAGTATCTCACTCT 1680
Db 614 CTTTCTCACTCCCAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 555
QY 1681 ACTCACTCCCTGACTTATGATGCTGAGACCAATTCAGTGTGAGAGTCAAGCTGA 1740
Db 554 ACTCACTCCCTGACTTATGATGCTGAGACCAATTCAGTGTGAGAGTCAAGCTGA 495
QY 1741 CGTCAAGTGTAGCCCAATTTGAGATTAATCAATCAAAAAAGTATCTCACTCT 1800
Db 494 CGTCAAGTGTAGCCCAATTTGAGATTAATCAATCAAAAAAGTATCTCACTCT 435
QY 1801 CCGGGAAGAGAGCTTGTGATCAATTTCTGCAAGCCCTGACTTATTTGCACTGA 1860
Db 434 CCGAAGAGAGAGAGCTTGTGATCAATTTCTGCAAGCCCTGACTTATTTGCACTGA 375
QY 1861 TGTATCAAGCAATATTTGCTTGTGCTGAGAGACAGACATTCAGTGTGAGAGG 1920
Db 374 TATTAATCAACATATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 315
QY 1921 TGGGGGTGTGTCTGCTGCACTCTAATATGAGAGTTCAATGCTGCAACCAAGCTGTC 1980
Db 314 TAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 255
QY 1981 TGTGGCTCTGAGCTCAGATGAGTGTAGGAGTGTAGGAGAGCCCTGCTGAGGTTG 2040
Db 254 TATTAATCTTAACACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 195
QY 2041 GAGAGCTATGTGTCTGAGCAATTCAGCCCTCTTCTCGTGTGAGTGAAGCTGGAG 2100
Db 194 AAACCTTAATTAATTTCTTAACATTAACCCCTCTTCTCGTGTGAATTAATTAATTA 135
QY 2101 TAGAAGCGAAGATTGTTCTGAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 2160
Db 134 TAAAAAGCAAAATTAATTTCTTAACATTAATTAATTAATTAATTAATTAATTAATTA 75
QY 2161 CAATTTTGGGGGTGAGGGGCTAATGAGAAAGGATCGCTCCATGATCAAGATATCTAT 2220
Db 74 CAATTTTAAATCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15

```

```

QY 2221 TTGTTCTCCCTA 2232
Db 14 TTATTTCTCCCTA 3

RESULT 7
HUMHXMA01
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Altitude, F., Pol, V., Restagno, G. and Silengo, L.
TITLE
Structure of the human hemopexin gene and evidence for
intron-mediated evolution
J. Mol. Evol. 27 (2), 102-108 (1988).
JOURNAL
MEDLINE
PUBMED
88316972
2842511
Original
source text: Homo sapiens DNA.
LOCATION/Qualifiers
1..731
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="11p15.5-p15.4"
178..288
/partial
/exon
/gene="HPX"
/note="G00-120-054"
/number=1
206..274
/gene="HPX"
/note="G00-120-054"
/number=2
289..438
/gene="HPX"
/note="G00-120-054"
/number=1
439..497
/gene="HPX"
/note="G00-120-054"
/number=2
498..617
/gene="HPX"
/note="G00-120-054"
/number=2
618..689
/gene="HPX"
/note="G00-120-054"
/number=3
690..731
/gene="HPX"
/note="G00-120-054"
/number=3
Intron
/gene="HPX"
/note="G00-120-054"
/number=3
BASE COUNT
148 a 167 c 204 g 212 t
ORIGIN
Query Match 5 8%; Score 681.4; DB 9; Length 731;
Best Local Similarity 98.1%; Pred. No. 5e-170; Indels 3; Gaps 2;
Matches 711; Conservative 0; Mismatches 11;
QY 1796 AGCTGCGGGAAGAGAGCTTGTGATTCAGATTTCTGTCCAGACCTGACTTATTTGC 1855
Db 1 AGCTGCGGGAAGAGAGCTTGTGATTCAGATTTCTGTCCAGACCTGACTTATTTGC 60
QY 1856 AGTATGTAATCAAGCAATATTTGGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1915

```

Db 61 AGTAGTATATACGACCAATATTTGGCTTAGTCTGGAGACAGACATTCCTCCAGTAGAGTT 120

QY 1916 GGAGGAGGGGGGCTGCTGCTGCAACTCTATATAGAGAGTTAACTGACACCCGAGAC 1975

Db 121 GGAGGAGGGGGGCTGCTGCTGCAACTCTATATAGAGAGTTAACTGACACCCGAGAC 180

QY 1976 TGTCTGTGGCCCTCTGACAGCTCAGCAATGGGTAGGTACTGGAGACACCCGTTGACATGG 2035

Db 181 TGTCTGTGGCCCTCTGACAGCTCAGCAATGGGTAGGTACTGGAGACACCCGTTGACATGG 240

QY 2036 GTTGTGAGGCTTATGCTGTCTGTGACCAATGGCCACCCCTCTTCTCCGTAGTAAAGCT 2095

Db 241 GTTGTGAGGCTTATGCTGTCTGTGACCAATGGCCACCCCTCTTCTCCGTAGTAAAGCT 300

QY 2096 GGAGCTAGAGAGCGAAGATTTAGTTTCTGGGCTAGGTTAGTTAGTTAGGCGCAATTTTAAAGCC 2155

Db 301 GGAGCTAGAGAGCGAAGATTTAGTTTCTGGGCTAGGTTAGTTAGTTAGGCGCAATTTTAAAGCC 360

QY 2156 TCGGTCAATTTTGGGCTCAGGGGCTATGGGAAAGGATCGGTCCCAATGGATCAAGTAT 2215

Db 361 TCGGTCAATTTTGGGCTCAGGGGCTATGGGAAAGGATCGGTCCCAATGGATCAAGTAT 420

QY 2216 CTATTTTGTCTCTCTAGAGCTAGTGGCCATGGGAATGTTCTGAAGGCGAGCAAGCC 2275

Db 421 CTATTTTGTCTCTCTAGAGCTAGTGGCCATGGGAATGTTCTGAAGGCGAGCAAGCC 480

QY 2276 AAGCCAGACGTACCTGTGAGAGCCCTGACTCCCTAAGTGTCTTATCTGTCTGTGTGT 2335

Db 481 AAGCCAGACGTACCTGTGAGAGCCCTGACTCCCTAAGTGTCTTATCTGTCTGTGTGT 539

QY 2336 GTCTCTGCAATTTTATGACCTTCTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2395

Db 540 GTCTCTGCAATTTTATGACCTTCTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 599

QY 2396 TCCCTAATCTTCCACCCAGAAAGCTGTCTGAGATGCTGAGCTTTGATGCTACCAACCTGG 2455

Db 600 TCCCTAATCTTCCACCCAGAAAGCTGTCTGAGATGCTGAGCTTTGATGCTACCAACCTGG 659

QY 2456 ATGACATATGGAACATGCTGTTTCTTTTAAAGTA--GAGGAGCTAGAGTTAGGCGCTTA 2513

Db 660 ATGACATATGGAACATGCTGTTTCTTTTAAAGTAAGTATGAGGAGGATTAAGGCTGCTTA 719

QY 2514 GGACC 2518

Db 720 GGACC 724

RESULT 8

HSHEMPRO

LOCUS HSHEMPRO 511 bp DNA linear PRI 16-MAR-1992

DEFINITION Human DNA for hemopoxin promoter.

ACCESSION X15399

VERSION X15399.1 GI:32060

KEYWORDS hemopoxin; promoter.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE Polt,V., Silengo,L., Altitude,F. and Cortese,R.

JOURNAL The analysis of the human hemopoxin promoter defines a new class of liver-specific genes

MEDLINE Nucleic Acids Res. 17 (22), 9351-9365 (1989)

PUBMED 90067933

REFERENCE

AUTHORS Polt,V.

TITLE Direct Submission

JOURNAL Submitted (11-JUL-1989) Polt V., EMBL, Meyerhofstr 1, D 6900 Heidelberg, F R G

FEATURES

source

1..511

/organism="Homo sapiens"

QY 1484 AAGCTTGAGAAAAGCTGTTCTCTGAGTATTTCTAAGTATTTATCTGTGTGTGTTA 1543

Db 1 AAGCTTGAGAAAAGCTGTTCTCTGAGTATTTCTAAGTATTTATCTGTGTGTGTTA 60

QY 1544 CTAAAGTATGATGATTTGCTGTCTGTAGCTGCTTTAGACAGGCTTTGACACATGAC 1603

Db 61 CTAAAGTATGATGATTTGCTGTCTGTAGCTGCTTTAGACAGGCTTTGACACATGAC 120

QY 1604 AGCAATATGATTTCCCTCTTTCTTCTGACCTCCCTCCCTAGGAGATTAATCAACA 1663

Db 121 AGCAATATGATTTCCCTCTTTCTTCTGACCTCCCTCCCTAGGAGATTAATCAACA 180

QY 1664 AAGGTATCTGATGATTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723

Db 181 AAGGTATCTGATGATTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

QY 1724 TGAAGTCAAGCTGAGCTCAGACGTGATGATGATGATGATGATGATGATGATGATGAT 1783

Db 241 TGAAGTCAAGCTGAGCTCAGACGTGATGATGATGATGATGATGATGATGATGATGAT 300

QY 1784 GGGTGGGGGGGAGCGCGCGGAGAAAAGAGATGATGATGATGATGATGATGATGATGATGAT 1843

Db 301 GGGTGGGGGGGAGCGCGCGGAGAAAAGAGATGATGATGATGATGATGATGATGATGATGAT 360

QY 1844 GACCTTATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1903

Db 361 GACCTTATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 1904 CCGAGTAGAGTTGAGAGTGGGGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1963

Db 421 CCGAGTAGAGTTGAGAGTGGGGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 1964 GTTACCCAGAGCTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1994

Db 481 GTTACCCAGAGCTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511

RESULT 9

AC125227

LOCUS AC125227 249769 bp DNA linear HTG 25-AUG-2002

DEFINITION Mus musculus chromosome UNK clone RP23-42113, WORKING DRAFT

ACCESSION AC125227

VERSION AC125227.3 GI:22476035

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE

AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE The sequence of Mus musculus clone

JOURNAL Unpublished

REFERENCE

AUTHORS McPherson,J.D. and Waterston,R.H.

TITLE Direct Submission

1..249769

McPherson,J.D. and Waterston,R.H.

QY	5752	TCCTACCACTTTCTCTCCAGGC--CATGACACAGGAATGGACCTGGCATTTGGGAC	5809
Db	127531	CCATGACCTGTTTCTCTCCAGGCATGGTAGCCGAGAAATGAACTGCTCATGGAAAT	127590
QY	5810	AGTACCCACATGGCCCTGATATATGCGCTGAGCCCATCTAGTCTTGTCTGCATG	5869
Db	127591	AGCACCCA-----TCCTATGATTCACGTTGATGACCAGATCTGGCCTGATGCATG	127644
QY	5870	ACGTCTGACACCAATGATGATGACCTTAGTGGCTTAGTGTGATGATGATGCCCACTCCC	5929
Db	127645	CTGTGTGACCAATCGGGGTGACACCTATGCTTCACGTGTGAGAGGGCTGACCTCAATCCC	127704
QY	5930	CAATGNGCTCTACACTCTTTTACTGTATATCCATCCCTGTGACACATTTCTCCATGT	5989
Db	127705	CATGTCTACCCCCCCCCACCCCGTCTCTTTTACTTATATGTCTGTCTTTAGACATAGCTT	127766
QY	5990	CATCACTGTGTCACTTATTTTGTCCCCCTGTGCCCCATCCCTCTGATGAGCCCTTGCAAT	6049
Db	127765	CTCCACTTTATCACACACATCCGATCATTTTGCCCTTGTATCTCTCTCCATATACAT	127822
QY	6050	CCCTCATCT-CTGAGGCATATTTCTCAATCTTGTCTGTCAAGGCCCAAG---CCCTTAA	6104
Db	127825	CCCTCATCTCCCTAGGTTCAATAGTCCAACTTATCTCTCATATCAGAGGCTTCTTTCCA	127888
QY	6105	CTTACATCAACGTGTACACCATCTACCTCCATGSGCTGTGCCCCCTGTGACCTCTGGGC	6164
Db	127885	TGCTTTTAACTGCTGTCTGCTCTTCAAGTATCTTACACCCCTGTCTGTATTTCTGTGGGA	127944
QY	6165	CCCTATGACTCCCTGTGTCTC---CTTGTCTAATGCCCTGTGAGAGCCCTCTGGCTCC	6221
Db	127945	CCCTCTTACCCATATGTAATTTCTGCAATGTTTCTCTGCTCTGTGATCTCTTACATGTC	128004
QY	6222	CTTGCTCCCTGGAACCTTATATGTCTCTGTACTCTTGTGCTCCCTTGTCTTTCATAT	6281
Db	128005	CTTGCTCTCTGAGGCTGTGTGTGTCCATG--CCGTGTGATCTCC--TGTGTTGTGGCAT	128061
QY	6282	CTTTCTGAGTCTCTGGGCTC-----	6301
Db	128062	CTCACTGAGACTCTCGGTTCTTGCTTGCTTCTGGAATCTCCCTGTGACTCTGTGCTCTTT	128121
QY	6302	-----CCCTGATTTATCTCTCA	6318
Db	128122	GTGTGATTTGTGTAAGACATGCGCTTTTGTGATATATTTGATTCATTAATTTATTTCA	128181
QY	6319	GAACTCCATCTTGTTCAGGTTCTGTGTTCTATGTCCAGACCCCTGGGCAATGACTGC	6378
Db	128182	GGAACTCCATTTTCCAAAGTACTTAGTCTTATCAATCCAGGTCTGTGGCTTGTGTAC	128241
QY	6379	CTGGGGATGAGATGTTCTCATTTGTGTGAGAACCAAGCTGAAAGATGTTGGTACTTTAGACC	6438
Db	128242	TTGGGGATGAGATGTTCTCAATGTCTTAAAGAAACGAAAGAAAC--TGAGTATTTCTAGACC	128299
QY	6439	TTTAAAGAGCTGGCTTCACTAGAGCTGTGAGGTTTCTCTCTGAGTGTGCCAATGGAATAC	6498
Db	128300	CTTAGAGTCTCCCTTCA-----AAGAGGTCCAGAGAAAGAA	128356
QY	6499	CCCTCCCTTGAACCCGTGCATCAATTTGTAAAGCCATCTAATTAATACCTAGGCTGT-T	6557
Db	128337	CCCAACTTGAACCTTTGTATCGGTTGTGTAAGCAATCTCCGATATCTAGAGACATAT	128396
QY	6558	CTGAGTTCAATGACGAGCATTAATATGATCATGCTGTGCACAGTTGAGATATCCCAAGAGAG	6617
Db	128397	AATCTCTAGGCTGAGGTTGACACATCTCTACACTGAGGACATCCCAAAATGGGGAG	128456
QY	6618	TGAGCAACCATCATCATCAACCTGAGATATATGATATAATTAAGCAGTGTGAAGATA	6677
Db	128457	GGGACATATCATATCACGTTTCACTCTGATGATACATGCA-----ACAAAAGTGGCTTAAATG	128511
QY	6678	TAAATATCGTGAATATATTTTCTCACACAAATTTTGTGGCTCTGACCTGTGACAA	6737
Db	128512	GCAAGGGGACAGAAAAT-----AGAAATACCTTTTATGTTCTGTAATGACTTAACCAA	128564
QY	6738	TTTGACCATTTATGACTATCAATTTCTGTGTGAATAATATCATCAACATGAGAGCAAT	6797

[illegible]

LOCUS	CNS01DM6	202495 bp	DNA	linear	PRI 28-APR-2001
DEFINITION	Human chromosome 14 DNA sequence BAC R-596D21 of library RPc1-11				
ACCESSION	from chromosome 14 of Homo sapiens (Human), complete sequence.				
VERSION	AL136418				
KEYWORDS	AL136418.4	GI:13897283			
SOURCE	HTG.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Mammalia; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 Helig, R., Pett, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P., Broclet, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F., Ley, M., Eckenberg, R., Bruls, T., deBeraudinis, V., Cruaud, C., Gysay, G., Saurin, W. and Weissbach, J.				
TITLE	Sequencing of the human chromosome 14				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 202495)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (26-APR-2001) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) On Apr 30, 2001 this sequence version replaced gi:12001732.				
	----- Genome Center				
	Center: Genoscope / Centre National de Sequencage				
	Center code: GS				
	Web site: http://www.genoscope.cns.fr/				
	Contact: SeqRef@genoscope.cns.fr				

	The following BAC sequence is oriented from the T7 to the SP6 end.				
	Upstream BAC (overlapping the T7 end) : R-176H8 (AC=AL139353)				
	Downstream BAC (overlapping the SP6 end) : C-2313013				
	----- Summary Statistics				
	Assembly program: Phrap, version 2.0				
	Quality coverage: 7.09x in Q20 Bases; sum-of-contigs				

	Overall quality chart :				
	Range	:	bases		
	0	:			
	1	:			
	10	:	9	:	1
	20	:	19	:	9
	30	:	39	:	112
	40	:	49	:	2205
	50	:	59	:	7368

```

60 - 69 : 11469
70 - 79 : 28289
80 - 89 : 66625
90 - 99 : 86417
-----
Percentage of bases with a quality value >= 40 : 99 %
Location/Qualifiers
1. 202495
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-596D21"
/clone_1ib="R-PCI-11"
18107..18281
/notes="matching EMBL:R22776
RHD:RH6015
dbSTS:STS45945
Identified using the e-PCR software (G. Schuler)"
23355..23600
/notes="matching EMBL:AA521333
RHD:RH98883
dbSTS:STS68632
Identified using the e-PCR software (G. Schuler)"
169462..169599
/notes="matching EMBL:R69806
RHD:RH53627
dbSTS:STS29238
Identified using the e-PCR software (G. Schuler)"
BASE COUNT 57490 a 41081 c 42091 g 61833 t
ORIGIN
Query Match 4.8%; Score 479.4; DB 9; Length 202495;
Best Local Similarity 54.9%; Pred. No. 4.6e-116;
Matches 1436; Conservative 0; Mismatches 1011; Indels 170; Gaps 18;

QY 7181 ATATTAAATATACCGAGCGCGCGGTGAGTCAAGCCCTGATATCTGACCTTTGGGA 7240
Db 97791 AAAAAAAAAAAGAAAGCCGCGCATGTTGGCTAATGCTGTAATCCAGACCTTTGGGA 97732
QY 7241 GGCTGAGCAGGTGATGCTTGAAGCGGCTTGAAGCTTGAAGCTTGAAGCAGCTTGGC 7300
Db 97731 GGCCGAGCGAGGTGATCAAC-----TGAAGTCAGAGAGTTCAAGACCACTTGGTC 97681
QY 7301 AACATGCAAAACCTCATCTCTCAAAAAATATATAAATTAATGTCGGGTGTAAGTCCT 7360
Db 97680 AACATGTAACCCCGTCTCTCAAAAAATATCAAGAAATTAATCGGCGGTGTAAGT 97621
QY 7361 TCCTATAGTCCATCTACTTCAAGAGGCTGAGCCGGA-----AGGTC 7402
Db 97620 GCTGTATATCCAGCTACTCGGAGGCTGAGGAGGAGAAATCACTTGAACTGAGAGTG 97561
QY 7403 GAGCTTCAGTGAAGCGGTATCGTGTACTGCACTCCAGCCTGGTGAAGAGAGCC 7462
Db 97560 GAGGTTCAGTGAAGCAAGTTGCAACACTGCACTCCAGCCTGGTGAAGAGAGCT 97501
QY 7463 ATGTCTCAAAAAAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACCC 7522
Db 97500 CCCTCTCAAAAAAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACCT 97441
QY 7523 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7582
Db 97440 CCCTTCAATATATATATATATATATATATATATATATATATATATATATATAT 97381
QY 7583 ATGACCCAGCTACAGATGGG-----CAGTCCCTCCCTCTCACTGGTAA 7628
Db 97380 GTCTCAGGCTTACGTGTTGTTCAGTCTGTAAGCAATTAAGCAATCGGTATATATAT 97321
QY 7629 ATTTTCTTCTCTGACTCAAGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 7688
Db 97320 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 97261
QY 7689 TCACTGTGACCCAGGCTGAGTGAAGTGGCAATCTTGGTTCACGTGAACTCTGCC 7748

```

```

Db 97260 TTGCTGTGCACTAGAGCTGAGTGAAGTGAAGCAATCTTGGCTCACTGAACTCGCC 97201
QY 7749 TCTGGGTCAACAGCATCTCTGCTCAAGCTCCGTAATAGTGGACTCAAGCGCAT 7808
Db 97200 TCTGTGCTGAGCCATCTCTGCTCAAGCTCCGTAATAGTGGACTCAAGCGCAT 97141
QY 7809 ACCAGCATGCGGCAATTTTGTATTTTGTGGTTACAAATGATATATATAT 7868
Db 97140 ACCAGCATGCGGCAATTTTGTATTTTGTGGTTACAAATGATATATATAT 97121
QY 7869 TTAATTTTGTATTTTGTATAGATAGATAGGATTTACATGTTGGCCAGGCTGTCGA 7928
Db 97120 ----TTTGTATATTTTGTATAGAGAGGAGTTTCAAGGTTTACAGAGTATGTCAT 97065
QY 7929 CTCTGACCTCAAGTATCCGCTGCTGCGCTCCCAAGTGTAGATTACAGCATG 7988
Db 97064 CTCTGACCTC--GTCAACCGCTGCTGCGCTCCCAAGTGTAGATTACAGCATG 97007
QY 7989 AGCAACGAGCGTGGCCCTCAATAGGTTTATATCTATCTTGTCTTCAACTT 8048
Db 97006 AGCACTGCGCCG--CGCATCACTACTTTAAGCTCTAGGTCAACATGTGA 96949
QY 8049 GGCTTCAGTGGACCATGATGTTCTCTCACTTCTCACTTCACTGATCTTCAATG 8108
Db 96948 GCGAAGCTGAGAACCACTGACTAGTCAATATATATATATATATATATATATATAT 96889
QY 8109 TCAGTTCACATGATACCTCCCTCACTGCTCTTTTCTTCTAGTAAGATTTCCAGAG 8168
Db 96888 TATCTGCTGCTGCTGCTGCTCCCAAGTGTGATTT-----ATGAGT 96846
QY 8169 GAATCTGAATGGCCCATGCTCATATTTTCAACCAACCATTAAGTGTATGCCA 8228
Db 96845 GTGAGCAGCTGACCCAGGTTGTTCTTATTAATAAATTTCGCAATGCGC 96786
QY 8229 GCTATGTAATGCTCATTAATAGGTTGGAACTCATCATTTCACTTCAATGCAAAAG 8288
Db 96785 TCTCGAATTCAGTATCATCAT-----TATATCTCATATACATTTCCG 96740
QY 8289 AGCATAGCTCTGTTCTCAAAATAGGCCCCGAGGCGAGGTGTGCTCATGCTTATA 8348
Db 96739 GGAAGCTGGTTAAAGGCAATATAGAAATTTTATAGGCAAGGCTCATGCTGTA 96680
QY 8349 ATCCCAACATGTTGGAGGCCAGAGGGGCAATCACTTGAATCCAGGTTTAAACCA 8408
Db 96679 ATCCCAACATTTGGAGGCCAGAGGGGCAATCACTTGAATCCAGGTTTAAACCA 96622
QY 8409 GCTGAGCAACATGTTGAATCTCATCTTCACTTCACTTCACTTCACTTCACTTCACT 8468
Db 96621 GCTGAGCAACATGTTGAATCTCATCTTCACTTCACTTCACTTCACTTCACTTCACT 96563
QY 8469 GTGGCATGACCAAGTATGCTCCAGCTGTGAGAGGCTGAGGAGATGTCAGATG 8528
Db 96562 GTGGCATGACCTGTATGCTCCAGCTTCTCGGAGGCTGAGGAGCAATTCGTTAACC 96503
QY 8529 TGGAGGCAAGATTTGATGAACTGATCTGCTCTGCAATCAAGCTGAGTGAAG 8588
Db 96502 CGGAGAGTGAAGTTGATGAGGCAAGATGCGCCCACTGCAAGCTTGGCCACAG 96443
QY 8589 ATTGAGACCTGCTCAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 8648
Db 96442 AGGAGATCTCATCTCAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 96383
QY 8649 GGTAAATTAAGTGGAAACCTCCAGGAGGTGATATGTCAGTCAAGCGGTATATCTC 8708
Db 96382 CTTTATTAACCCCAAACTAATTTCAAAATAAATAAATAAATAAATAAATAAATAAATA 96323
QY 8709 AGTACAGGCTTAATAGAACTTGTGTGAGCAAGCAACACTAGATATTAATCAACA 8768
Db 96322 AAGTGACTAATCTTATGATATGCTGCTTCAAGATCACTGCTTAATTTTCAAAAGATT 96263
QY 8769 AATATTTGTTGACATCTGATTAAGAGTGGGCAATGCTTACAGCACTGATACATAGT 8828

```



```

Db      96262 CAAAAATGATTTATTTAAAAATTAATTCGAAATTTG-----GACAAATTTGTTACATGATT 96208
Qy      8829 CAACATGGCAGACAAGATGCTGCGCTGACAGAGCTCTGTAATAGTGAGAGACATATA 8888
Db      96207 CACTATATCTTACATATCTCCCT-----TGGTAGTGAAGAAAAA 96168
Qy      8889 GAAAGAGAAAGAAAGAAAGAAATTAATTTTAGTATATTAAGGGTGTAAAGAAATAA 8948
Db      96167 AAATATATATGTTAAATGTTGCTTTATGTTTCTAGTATGTTTCCATGGCTAAA 96108
Qy      8949 GACAGATAGTGGATAGAGTGAAGAAATGAGGCTGCTCTCTGAAGAAATGATTTT 9008
Db      96107 AAATCTTTATATATATGCTCTTTGTAAACATCTGTCCTCTTATCTTAAGGAAATTTA 96048
Qy      9009 GACCTGAGCTTACATGATGAGAAAGAAATTAACACAGCATGTCGCTGAGAGAAAGCATT 9068
Db      96047 TCTTTCGCAATTAATTAATATATGTTATTAATCTGATTAACCTTAATCTGTATACATAATAT 95988
Qy      9069 TTAGGAGGGTGAGCAGACATATCTTCAGAGAAATCAAGAAAGAAAGCTGCTGGA 9128
Db      95987 TACTGTCATAAACATATCTTGGATACATTAATTAATTAATTAAGCCAGTTAAAGAAA 95928
Qy      9129 ACACAGAAAGAGCAGGTGGTGACTTGAAGGGCAGGACGGCAGTGGCCAGTTACC 9188
Db      95927 AGTGGATCTAATTTGTAAATATGATATCAATAGAA-----ACTTCAAGAACCCA 95876
Qy      9189 TAGACCTGTAGAGGTTTCAACCATTAAGGGAGTACAGAAAGCTTGAGAGGGCT 9248
Db      95875 GGGCCAGGCAAGATGCTCTCTCTGTAATCTAGCATTTGGTAGGCTAGAGAGAGG 95816
Qy      9249 GTGATATATTTCTAATCTATTTTATTAAGATCACTCTGACTTTTTCAGAAATATAGT 9308
Db      95815 GTGATCACTTGAAGCCAGAGATTAAGGACACAGCTGGGCAAGGTAGAGAGACCCATCT 95756
Qy      9309 TATTAAGTACAGCATATGACAGAGAAATCCAGCTAGCAATCCGTCAGTTTCCAAAT 9368
Db      95755 CTTTAAGAAAAATTAATTTTATTTTAAATC-----TAAAGTTTTTTTAAAGA 95703
Qy      9369 TAGAGGATGACCGCTGAGTACAGATGATGATGACAGAGAGGTGAGAGATCACATG 9428
Db      95702 AAATCTTCAAGAACCTTAATTTCTGTAGATATATTAATTTTCTTAATTAAGATTCGAAAT 95643
Qy      9429 ATATATTTTGAAGTGAAGCTGACAGCATTAATTAATAGCTTAAGATAGCCGGG--TGTG 9486
Db      95642 GTGACATTAATGCGGAAAAAAGTATATTAATTTGTTGTAATGATTTTGAAGAACGGGCCAGGG 95583
Qy      9487 GTGGCTTAAGCTCTGTATCTTACGACTTTGGAGGCCA---AGCCAGTGGATCACTGA 9543
Db      95582 GTGGCTCATGCTGTATATCCAGACACTGTGGAGGCCACCGAGGTGGTGGATTAACCTGA 95523
Qy      9544 GGTCAAGAGTTTGAAGACAGCTTGAACCAATGATGTAACCTGCTCTAATAAAATAC 9603
Db      95522 GGTCAAGAGTTTGAAGACAGCTTGAACCAATGATGTAACCTGCTCTAATAAAATAC 95463
Qy      9604 AAAATTAAGCTGGAAATGTTGCAATGCTGTATCTCAGCTTAATCTTGGAGGCTGAGGC 9663
Db      95462 AAAATTAAGCTGGGCGGTGCTGCTAATGCTGTATCTCAG--CTACTCAGAGGCTGAGGC 95404
Qy      9664 AGAGAAATGCTTGAAGCTTGGAGGTGAATGTTGCAAGTGGCCGAGATTTGCACATTTGA 9723
Db      95403 AGAGAAATGCTTGAAGCTTGGAGGTGAATGTTGCAAGTGGCCGAGATTTGTGCACATTTGA 95344
Qy      9724 CTCACGCTGGGGAACAAGATGAAATCGCTCTTA 9760
Db      95343 CTCACGCTGGGGAACA-AGAGCAAGACTCATCTCAA 95308

```

```

RESULT 11
AC108045/c
LOCUS      AC108045
DEFINITION Homo sapiens BAC clone RP11-307L18 from 4, complete sequence.
ACCESSION AC108045 AC021616
VERSION    AC108045.3 GI:27356779

```

KEYWORDS SOURCE ORGANISM

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
9847074
2 (bases 1 to 74904)
Armerstrong,J., Haglund,K., Courtney,L., Mangiapanello,L., Trani,L.
and Bielicki,L.
The sequence of Homo sapiens BAC clone RP11-307L18
Unpublished (2001)
3 (bases 1 to 74904)
Waterston,R.H.
Submitted (24-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 74904)
Waterston,R.H.
Submitted (23-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 74904)
Waterston,R.H.
Submitted (21-DEC-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 74904)
Waterston,R.H.
Submitted (27-JAN-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Dec 21, 2002 this sequence version replaced gi:18873910.

REFERENCE AUTHORS TITLE JOURNAL

REFERENCE AUTHORS TITLE JOURNAL

REFERENCE AUTHORS TITLE JOURNAL COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
Summary Statistics
Center project name: H_NH0307L18
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frenken,E., Tateno,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved

Db	2616	ACCTGGAGGCTGGAGGTTGACGTGAGCTTAATTGCACTGACCTCAGCTGGGGGA	2557
Oy	7451	CAGGTGAGACCATGTCTCAAAAAAAAAACAAAAACAAAAACAAACAACA	7510
Db	2556	CAGAGCTAGACTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGA	2497
Oy	7511	AACAAAAAACCATATATATATATATATATACCTAGCTGAGTGAGAAATCATATTTTGGT	7570
Db	2496	ACATATTAACGAGGAAAT-----TACAAATCAACTTTTTTTTTTTG	2452
Oy	7571	AAAAATCAACAATGACCCAGCTACAGATGGGGAGTCCCTCCCTCTCACTGTAAT	7630
Db	2451	AGACAGAGTCTCTTATATTTCTCTCACTGCTGTTGCTGTGGTAAGATGAGTATAT	2392
Oy	7631	TTTTCTTCTCTGACTCAAGTTTTGTGTGTGTGTGTGTGTGTGTGATGAGTCTC	7690
Db	2391	TTCTTTCTTTCTTTCTTTTTTTTTTTTTTTTTTTTTTTTGAAGCGAATCTTCTC	2332
Oy	7691	ACTGTGACCCAGGCTGAGTGCATAATGGGCAATCTTGTTCACTGCACTGCTGCTC	7750
Db	2331	ACTGTGCCCCAGGCTGAGTGCATGAGGCGAATCTTGCTCACTGCACTTCCCTCTC	2272
Oy	7751	CTGGGTTCAAGCGATCTCTCTGCTCAGCCTCCCGTATAGCTGGGACTACAGGCGATAC	7810
Db	2271	CGGGGTTCAAGCGATCTCTGCTCAGCCTCCCAAGAGCTGGGATTAAGAGGGGCTGCTC	2212
Oy	7811	CACCATGCTGCTGCTAATTTTTTGTATTTTTTTTGGGTTCAATGACTATTATTAATTT	7870
Db	2211	CACCATGCTGCTGCTAATTT-----	2194
Oy	7871	AATTTTGTATTTTATGAGATATAGGTTTACCATGTGGCAGGCTGCTCGAAT	7930
Db	2193	-----TTTTGTATTTTGAAGAGACAGGGTTTCCATTTTGAAGGCTGGTCTCAACT	2136
Oy	7931	CTGACCTCAGGTGATCCGCTGCTCGGCTCGGCTCCCAAGTCTGAGATTAAGGCAATGAG	7990
Db	2135	CTGACCTCAAGTATATCGGCCCACTGGGCTCCCAAGTCTCAGATTAAGGCAATGAG	2076
Oy	7991	CAACCAAGCTCTGGGCCCCCTCATAGGTTTATATCTATTTCTCTTGCTCTTCAACCTTTGG	8050
Db	2075	CCAGGCGCCCAAGCCAAAGTATATCTTTCAAGGTAATCTGTCACTTAAATGTATA	2016
Oy	8051	CTTGAGGTGAGCAATCATATGTTCTCTCCACTTCTCACTACTATGATCTTTCAAGTCTC	8110
Db	2015	TAACTTTAATTAAGAATATTCATTTTCAGAAATATCTAGAAATCAAGAAAGTACA	1956
Oy	8111	AGTTCCAAGTATATCTCCCTCAGTGTCTTTTTTCTTATGAAATTTTCAGAGAGGGA	8170
Db	1955	AAGATGTCATATATGATGATTTCAAGTGCACTTTGTATATATCTATTTAAACGTCTCT	1896
Oy	8171	ATCTGAATGGCCCAAGTCATATTTTCAGACCAACCACTTAAAGTGTTGATTGCCAGC	8230
Db	1895	ACGTAA-----TAAATCAAGAAATCTATATTAACTTTGAATACGA	1853
Oy	8231	CTATGTAATGGCTATTAATAGGTTGGGAATCATCTATTACTTCATGTGACAAAGAG	8290
Db	1852	CAAGGCTGATATATGATGATGAGAAAGAGTCCACAGGTTTAAATATATATTTATA	1793
Oy	8291	CATAGCTCTGTTCTCAAAATAGGGCCCTGAGGCGAGGTTGTTGGCTCATGCTATAT	8350
Db	1792	TATACA-----GGCCAGGCTCAGTGGCTCACACTGTAT	1758
Oy	8351	CCCAACA-CTGTGGAGAGCCGAGGGGGGAGATCACTTGAATCCAGAGTTCTAGACAG	8409
Db	1757	CCAAACAGCTTTGGGAGGCGCAAGGAGCAATCTCTTGAAGCCAGAGTTCAAGACAG	1698
Oy	8410	CTTGGGCAACATGTTGAAATCTCATCTTACTTAAATAACAAAAATTTAGCAGGTGTGG	8469
Db	1697	CTTGGCCCAAAATGACGAACCCCACTCTCACTAAAAAAAC-AAAAATTAAACAGTGTGG	1639
Oy	8470	TGGCATGCAACAGTATCCAGCTGTTTCAAGAGGCTGAGGTGGAGGATTTGCTGAGTCT	8529
Db	1638	TGGGCATGCTGTATATCCACGCTACTTGAAGGCTGAGGAGGAGATCACTTGAATCC	1579

OY	8530	GGAGGCGAGATTGCAAGTAAACCGTACCTGTGCAATCCAGCTGGGGACAGA	8589
DB	1578	AGGAGGTGAAGTTGCAAGTAGCCGAGACCTGCACACTGCACTCCAGCTGGGGCAACGA	1519
OY	8590	TTGAGACCCGTCTCAAAAAACAATTAATAAATAAATAATATGTTCCTGAGCAGG	8649
DB	1518	GTGAGACTGTGATTAAAAAAAAAAATTAATATAC-----GT	1480
OY	8650	GTAATTTCAGTGGAAACCTCCAGGGAGGTGATATGTCAGTCACCGCTGATACTCA	8709
DB	1479	ACAGATTACAAGTAAAGAAAGCCCTAAGAAAGAAATTTTTCATTTCTCATGTTTTACA	1420
OY	8710	GTACACGGCTTAATAGAGACTGTGTGATACAGACAAACAATAGATTTATCTCAACA	8769
DB	1419	TGAGAGAAATATCTGTCATGTCTATAGAGAACCTTAAGCTGAAATTAACAGCCGTT	1360
OY	8770	ATATTGTTGAGCATCTGATTAAGAGTGGGCAATTGCTTAGGCACTAGATTAACGATGTC	8829
DB	1359	GAGAAATTTGTTTAATTTTAAAAATTTAAACAATCTCATTTTTCAAACATCATTAG	1300
OY	8830	AACATGGCAACAAGATGCTGCTGCAAGGCTCTGCTAAAGTAGAGAGACATATAG	8889
DB	1299	TAAAGACAGACTATATATACAGTGCATCTTTTTGACCAATTAATTCATCTTTAAAT	1240
OY	8890	AAAGAGAAAGAAAGAAAGAAATTTTATAGTATTTAAAGGTTGTAAAGAAATAG	8949
DB	1239	GTTATCTTAGGCGGCGGACAGTGGCTCAGCCTATATATCCACACTTTGGAGGCCAA	1180
OY	8950	ACAGATATGTGGGATAGAGGTGAGAGAAATGAGGGCTGTCTTCTGAAGAATGATTTTGG	9009
DB	1179	CGCAGCGAGATTACTTAGAGTCAAGAGTTTGAAGACAGCTGGCCAACTGGCAAAACC	1120
OY	9010	AGCTGAGACTTCA----GTATGAGAAAGAAATTAACCAACGATGTGCTGAGGAAAGC	9065
DB	1119	CGTCTACTATAAAAAACAAAAATTTAGCCAGGCAATGACCGGGCTCTGTTATCTTAG	1060
OY	9066	ATTTTAGGAGAGGTGACAGACACTACTTCAAGGAATCAAGAAAGAAAGCCTGTGAGGCT	9125
DB	1059	CTACTCTGAGGCTGAGGACGAGAAATCGGTTGAACCCAGAGGTTAAGGTTGCAAGTAG	1000
OY	9126	GGAACACAGAGAAAGACAGGTGGGTGACTTGAAGGGCAGGGACGGC-----AGTGGC	9179
DB	999	CCGAGATCGACCAACCACTCCAGCTGAGAGCAGAGGAGAGTCTCATCTCAATATAA	940
OY	9180	CAGGTTACTAGACCTGTGAAGGTTTTCACCAATAAAGAGATCATCAGAAAGCTTG	9239
DB	939	TAAATTAAGTAAATAATTAATTAATTAATAAACGTTATCTTTAAACCTGGACCTTAACG	880
OY	9240	AGCAGGGCTGTGATATATTTCTAATACTTTTATTAAGAATCACTGTGACTTTTGGAC	9289
DB	879	GAAATTAACGTAAATAGTTAAAGAAATAATATATTTTAAAGGCTAGAGTACTTTTAAC	820
OY	9300	AACATAAGTTATAAAGTACAGAGATGTAAAGCAAGAAATCAGCTAGCAATCCGTGCAGT	9359
DB	819	AAAACTACACACAGACAGACCTATTTCACTACGCTTAATTTCTTAAGACATATATTTT	760
OY	9360	TGTCCTCAATTAAGAGTATGACCGCTTGGACTAGATATAGACAGCAGAGTGGTGAAGA	9419
DB	759	TGGCAAAATTTTAATTAAGAAATATGCACTGATGATATCAATATTTTCACTCT-----T	707
OY	9420	ATCACCATGATATATTTTGAAGGTAGAGCTGACGACTTAATCTATATATGCTAAGATAGCC	9479
DB	706	ATCATATTTTAAAGGAATATGTGGTAAAAATTAGTGCTTTTAAAAATGTGAAGAAATGGCC	647
OY	9480	GGGTGTGTGGCTTAACCCCTGTATCTCTACACTTTTGGAGGCCAAGGCGAGTGTATAC	9539
DB	646	AGGGGCGGTGCTATCTCTGTATATCCACACTTTGGAGGCTGAGGACAGGTGTATCAC	587
OY	9540	CTGAGGTCAAGATTTCCAGACAGAGCTTGACCAACATGTGTAAACCTGTCTTAATATAA	9599
DB	586	CTGAGGTCAAGAGTTCAAGACACAGCTGGCCAAATGTATTAACCCTCATCTTAATATAA	527

QY 9600 ATA-CAAAATTAGCTGGGAATGGTGGACATGCTGTAATCTCAGCTTACTGGGAGGCT 9658
 DB 526 ATTTAAAAAATACTGGCGCTGGTGGTGAACGCTTGTATCCAG-CTACTGGGAGGCT 468
 QY 9659 GAGGACGAGGAATGCTTGAACCTGGAGGTGAATGTTGACGTAGCCGAGATTGACCA 9718
 DB 467 GAGGACGAGGAATGCTTGAACCTGGAGGTGAATGTTGACGTAGCCGAGATTGACCA 408
 QY 9719 TTGCACTCCGAGCTGGGGAACAGAGTGAATCCGCTCTAATAATGAATGAA 9774
 DB 407 CTGCGCTCCAGGCTTGCGAC-AGAGTGAAGACTGTCTCAAAAAA 353

 RESULT 12
 CNS10DX6/c 202496 bp DNA linear PRI 09-FEB-2000
 LOCUS
 DEFINITION Human chromosome 14 DNA sequence *** IN PROGRESS *** BAC R-596D21
 of RPCI-11 library from chromosome 14 of Homo sapiens (Human),
 complete sequence.
 AL139054
 ACCESSION AL139054.1 GI:6967150
 VERSION HTG; HTGS-DRAFT.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 202496)
 REFERENCE
 AUTHORS Direct Submission
 TITLE Submitted (09-FEB-2000) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT
 IMPORTANT: This sequence is unfinished and does not necessarily
 represent the correct sequence. Work on the sequence is in progress
 and the release of this data is based on the understanding that the
 sequence may change as work continue. The sequence may be
 contaminated with foreign sequence from E.coli, Yeast, vector,
 phage, etc.
 FEATURES
 source
 location/Qualifiers
 1..202496
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone="R-596D21"
 /clone_11b="RPCI-11"
 /clone_11c="41093 g 61811 c 10 others
 BASE COUNT 57474 a 41093 c 42108 g 61811 t 10 others
 ORIGIN
 Query Match 4.7%; Score 467.4; DB 9; Length 202496;
 Best Local Similarity 54.8%; Pred. No. 7.3e-113; Indels 171; Gaps 19;
 Matches 1435; Conservative 0; Mismatches 1011;

QY 7463 ATGCTCAAAAAAACAACAAAAACAACAAAAACAACAAAAACAACAAAAACAAC 7522
 DB 97504 CGCTCTCAAAAAAACAACAAAAACAACAAAAACAACAAAAACAACAAAAACAAC 97445
 QY 7523 ATATATATATATATATATATATATATATATATATATATATATATATATAT 7582
 DB 97444 CCTTCACAAATATATATATATATATATATATATATATATATATATATATAT 97385
 QY 7583 ATGACCAAGCTACAGATGAGG-----CAGTCCCTCCCTCTCACTGTTAA 7628
 DB 97384 GTCTCAGGCTTACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 97325
 QY 7629 ATTTTCTTCTCTCACTCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7688
 DB 97324 ATATATATATATATATATATATATATATATATATATATATATATATATAT 97265
 QY 7689 TCACCTGTCAACCAAGCTGAGTGAATGAGCGCAATCTGTGTTCACTGCACTGACC 7748
 DB 97264 TTGCTGTGTCACTAAGCTGAGTGAATGAGCGCAATCTGTGTTCACTGCACTGACC 97206
 QY 7749 TCTGAGTTCAAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7808
 DB 97205 TCTGAGTTCAAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 97146
 QY 7809 ACCACATGCTGAGCTAATTTTGTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 7868
 DB 97145 ACCACATGCTGAGCTAATTTTGTATTTTGTGTTGTTGTTGTTGTTGTTGTT 97126
 QY 7869 TTAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTAT 7928
 DB 97125 ----TTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTAT 97070
 QY 7929 CTCTGACCTCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7988
 DB 97069 CTCTGACCTC--GTGAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 97012
 QY 7989 AGCAACAGCGCTGCGCCCTCAAGTTTATCTATCTCTGTTGTTGTTGTTGTTGTT 8048
 DB 97011 AGCAACAGCGCTGCGCCCTCAAGTTTATCTATCTCTGTTGTTGTTGTTGTTGTT 96954
 QY 8049 GGTTCAGTGAAGCAATCATGTTCTCTCACTTCTCACTTCTCACTTCTCACTTCTCA 8108
 DB 96953 GCAAGCTGGAAGCAACAGCTGAGCTGAGTGAATATATATATATATATATATATAT 96894
 QY 8109 TCAATTCCACTGATACCTCTCTCACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8168
 DB 96893 TCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 96851
 QY 8169 GAATCTGATGAGCCAGTCCATATTTTCAACACACACATTAAGTGTGATGCCA 8228
 DB 96850 GTGAGCACTGACCCAGCTGCTCTCTATTAATAAATTAAAGAAATCTTGCCAAATGCC 96791
 QY 8229 GCTATGATTTGCTCATTAATGAGTGGGAATCATCATTTCACTTCACTTCACTTCA 8288
 DB 96790 TCTCAAAATCATGATCATAT-----TATATCATCTCAATAATCATTCGG 96745
 QY 8289 AGCATAGCTCTGCTCTCAAAATAGGCCCCCTGGGCGAGGTGTGTGCTCATGCTTAT 8348
 DB 96744 GGAAGCTGGTTAAGGCAATATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 96685
 QY 8349 ATCCCAACATGTTGAGGAGCCAGAGGAGCAATCATTTGAGTCCGAGATTTTACCA 8408
 DB 96684 ATCCCAACATGTTGAGGAGCCAGAGGAGCAATCATTTGAGTCCGAGATTTTACCA 96627
 QY 8409 GCTGGGCAACATGTTGAGTATCTATCTTCACTTCACTTCACTTCACTTCACTTCA 8468
 DB 96626 GCTGACGATATATATATATATATATATATATATATATATATATATATATATATAT 96568
 QY 8469 GTGGCATGACATATATATATATATATATATATATATATATATATATATATATATAT 8528
 DB 96567 GTGGCATGACATATATATATATATATATATATATATATATATATATATATATATAT 96508
 QY 8529 TGGGAGGCAAGATTTGATGATGAAAGTGTGCTGTGCAATCAAGCTGGGTGACAG 8588

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Oosagawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6**NEIGHBORING SEQUENCE INFORMATION:**

Actual start of this clone is at base position 1 of RP11-163017; actual end is at base position 187360 of RP11-163017.

The region between bases 32935-33083 is single stranded and the sequence fidelity can not be guaranteed. RP11-163017 contains an unresolved dinucleotide repeat between bases 47950-48140 where the sequence fidelity can not be guaranteed. Digest information suggests that approximately 88 bases are missing from the current assembly.

FEATURES**source**

1.187360 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-163017"
/clone_lib="RPCT-11"
1.85
/note="similar to EST AA397735 (NID:g2050626) zc86a05.r1"
1.85
/note="similar to EST A1178109 (NID:g3728747)"
1.85
/note="similar to EST AUI20774 (NID:g10936009)"
1.85
/note="similar to EST AV610323 (NID:g9745993)"
1.85
/note="similar to EST BB074520 (NID:g8422511)"
1.85
/note="similar to EST BB614739 (NID:g9896336)"
1.85
/note="similar to EST BE786771 (NID:g10207969)"
209.800
/note="similar to EST BG432741 (NID:g13339247)"
227.725
/note="similar to EST BP911522 (NID:g12302980)"
649.800
/note="similar to EST AA397735 (NID:g2050626) zc86a05.r1"
649.800
/note="similar to EST A1178109 (NID:g3728747)"
649.800
/note="similar to EST AUI20774 (NID:g10936009)"
649.800
/note="similar to EST AV610323 (NID:g9745993)"
649.800
/note="similar to EST BE614739 (NID:g9896336)"
649.778
/note="similar to EST BE786771 (NID:g10207969)"
649.777
/note="similar to EST BB074520 (NID:g8422511)"
651.800
/note="similar to EST BG703239 (NID:g133975373)"
655.800
/note="similar to EST T32984 (NID:g615082)"
1023.1310
/rpt_family="Alu"
1309.1334
repeat_region

/rpt_family="(TTAA)n"
1678.1878
/rpt_family="MIR"
1938.1965
repeat_region
/rpt_family="AT_rich"
2379.2416
/rpt_family="AT_rich"
2399.2711
/rpt_family="Alu"
3034.3091
/rpt_family="MIR"
3295.3779
/rpt_family="L1"
3784.3834
/rpt_family="L1"
3913.4049
/rpt_family="MIR"
4261.4340
/note="similar to EST A1178109 (NID:g3728747)"
4261.4340
/note="similar to EST AUI20774 (NID:g10936009)"
4261.4340
/note="similar to EST AV610323 (NID:g9745993)"
4261.4340
/note="similar to EST BE614739 (NID:g9896336)"
4261.4340
/note="similar to EST BG432741 (NID:g13339247)"
4261.4340
/note="similar to EST BG703239 (NID:g133975373)"
4261.4340
/note="similar to EST T32984 (NID:g615082)"
4261.4318
/note="similar to EST AA397735 (NID:g2050626) zc86a05.r1"
5010.5100
/rpt_family="L2"
5163.5469
/rpt_family="Alu"
5165.5191
/rpt_family="(TTG)n"
5503.5802
/rpt_family="Alu"
6252.6283
/rpt_family="AT_rich"
6420.6447
/rpt_family="(T)n"
6840.6916
/rpt_family="MIR"
6917.7061
/rpt_family="Alu"
7062.7376
/rpt_family="Alu"
7347.7388
/rpt_family="AT_rich"
7377.7412
/rpt_family="Alu"
7413.7448
/rpt_family="Alu"
7434.7465
/rpt_family="AT_rich"
7802.8097
/rpt_family="Alu"
8200.8459
/rpt_family="Alu"
8604.8915
/rpt_family="Alu"
8886.8917
/rpt_family="(A)n"

Query Match 4.4%; Score 444.8; DB 9; Length 187360;
Best Local Similarity 53.5%; Pred. No. 7.7e-107;
Matches 1431; Conservative 0; Mismatches 1122; Indels 123; Gaps 19;

QY 7158 CCACGACATGAGAGACCTTATATTAATAATCCACGCGCGCGTGTGATGCACG 7217

Db	54129	CCATGCCCTCAGTGTGACCTTGATTGAACAAACAGGAGGCCAGAGTAGTAGGCTGACG	54188
Qy	7218	CTGTAAATCTTAGCACTTTGGAGGCTGAGGACAGTGGATTGCTTGAAGCGCTTGAAGCTT	7277
Db	54189	CTGTAAATCTTAGCACTTTGGAGGACCGAGGCGGGTGGGTGAT-----CACGAGTTC	54241
Qy	7278	AGGAGTTTGAGACCAAGCTTGGGCAACATGGCAAAACTCATCTCTACAAAAATA-TAA	7336
Db	54242	AGGACAGTGTAGACCAAGCTTGGCTTAACTGTGTAAACCTGGCCCTTATCTAAAAATACAAA	54301
Qy	7337	AATTAGTCGGGTGTGGTAGTGCCTTCTTAATGTCCTCATCTACTTCTGAGAGC-----	7387
Db	54302	AATTAGTCGGGTGTGGTAGTGCCTTCTTAATGTCCTCATCTACTTCTGAGAGGCAAGGCAAG	54361
Qy	7388	-----TGAGCCCGGAAGGTGAGGCTTCAATGAGCCGTGATGTCCTACTGCACTC	7438
Db	54362	AGATTGCTTGAACCTCTGAGGCAAGGTTTGAGTAGTGAATGTGTGCACACTACATC	54421
Qy	7439	CAGCTTGGGTGACAGAGTAGACCATGTCTCAAAAAAACAAAAACAAAAACAAAAACA	7498
Db	54422	CAGCTTGGGTGACAGAGTAGACCATGTCTCAAAAAAACAAAAAACAAAAAACAAAAAGAT	54481
Qy	7499	AACAAACAACAAACAAAAAACCATATATATATATATATATATATATATATATATATATAT	7558
Db	54482	AA-AGATGTGTAAATCAATAAAGACAGTCATATATCTGTCTGTATTTATATAG-CTATC	54539
Qy	7559	CACATATTTGGTAAATCAACAAACATGACCCAGCTACAGCATGGGAGAGTCCCTCCCTC	7618
Db	54540	CTGGCTCTGGAATATGTCTCTGTCTATTAATATATATATATATATATATATATATATAT	54599
Qy	7619	TCACTGTAAATTTTCTTCTCTGACTCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGT	7677
Db	54600	TTAAAAAAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT	54659
Qy	7678	-----GAAATGAGTCTCACTCTGTGACCCAGGCTGAGGTCATATGGCGAAATCTTGG	7730
Db	54660	CCCCCAAGACAGAACTTAAACCAAGTACCCAGGCAAGAGTCAATGGCACTATATAC	54719
Qy	7731	TTCACTGCAACCTCTGCTCTCTGGGTCAAGCATCTCTCTGCTGAGCTCCCGTATAG	7790
Db	54720	CTCACTGCACTTCAACTCTGAGATTAACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	54779
Qy	7791	CTGGGACTAAGCGGCAACCAACCATGCTGCTGCTAATTTTGTATTTTGTGGGTAC	7850
Db	54780	CTGGGACTTATGACATGCCCAACCAAGCTGCTGCTAATTTTAAATTTATTTT-----	54832
Qy	7851	AATCTACTATTTATTAATTTATTTTGTATTTTGTATGATATAGGTTTACACATGTT	7910
Db	54833	-----TTTGTAGAAAGAGGTCTTAAACATGTTGT	54862
Qy	7911	GGCCAGGCTGATCTCGAACTCTGACCTCGAGTATCCGCTGCTGCTGCTGCCCAAGT	7970
Db	54863	GCCAGGCTGATCTCGAACTCTGACCTTGAAGATCTCTTGGCTTGGCTCCCAAGT	54922
Qy	7971	GCTAGATTAAGGCACTGACCAACAGCTGAGCCCTCCTCATAGTTTTATCTATCTCT	8030
Db	54923	GTTGGATTAAGGCTGAGCACTGAACCTGGCCGAGACC-----TTTTCTGTATGAAG	54978
Qy	8031	TTGCTTTCTCAACATTTGGCTTGCAGTGGACATCATATTTCTCTCACTTTCTCACTA	8090
Db	54979	TTTGGACTTTGCTCTGTATTTCAAGGAGATGCCACTCTATATATTTCTAGCTCTTTT	55038
Qy	8091	CTTATGATCTTTCAAGTCTGAGTTCAACATATACCTCCCTGAGTTGCTTTTTCCTA	8150
Db	55039	AACAGAACAGTAAAAAGTATGTTCTTCTGACAAAGAACTTCAAAATATATAGAACAA	55098
Qy	8151	GTAAGATTTCCAGAGGGAATCTGAATGGCCAGTCCATATTTTGAACAACAACACAT	8210
Db	55099	AAATGTGCAAGAGGAGAGATTAACAAATCTCAACACATATTTTGGGAAATTTTAAACCC	55158
Qy	8211	TAAAGTGTGATTTCCAGCTTATGATTTGGCTACATTAATGGGTTGGGAATCATATT	8270

Db	55159	TAATAATTTATTAAGATATTCAAAGATCTGAATTACTGATCAACAGCTTTTGGTAAC	55218
Qy	8271	TACTCATTTGCACAAGACAGCATAGCTCTGGTTCTCAAAATAGGGCCCTGGGCCAGGTG	8330
Db	55219	GACATTAATAGTAATACCTACCCCTCAGC---ACATTAATTAATAAAAAATGTTCTTGGCTGGCCA	55275
Qy	8331	TGTGGCTCATGCTCTTAATATCCCAACACTGTGGGAGGCCGAGGGGGGAGATCATTTGAG	8390
Db	55276	AGGGGTCACTGCTCTTAATCCGACACTTCGGAGGCCAAGAGGGGGGAGTTGCTTAAG	55335
Qy	8391	TCCAGAGTTCTAGACCAAGCTTGGGCACATGGTGAAATCTCATCTTAATAAATATAC	8449
Db	55336	GTCAGAGGTTCCAGACCAAGCTGGGCCAATGGTGAAACCTGTCTTAATAAATATCA	55395
Qy	8450	AAAAAATTACCAAGGTGTGGTCATGCACAGTAGTCCACGCTTTAAGAGGCTGAGG	8509
Db	55396	AAAAAATTAACTGGGGGTGGTGCATGCACCTGTATCCCACTACTTGGGGGCTGAGG	55455
Qy	8510	TGGAGAGATTGCTTCAGATGTGGGAGGCAGAAATTCAGTGAACCGTAGCTGGCTCGC	8569
Db	55456	CAGGGGAATGCTTGAACCGGGAGGTGAGGTGACAGAACCCAGATTTGGCCATGCG	55515
Qy	8570	AATCCAGCCTGGGTGACAGATTGACACCTGTCTCAAAAAACAATTAATAAATAAT	8629
Db	55516	ACTCCAGCATGGGTGACAGGGCAAGCTGTCTCAAAAAAATAATTTCTTTAAAT	55575
Qy	8630	AAATATGTTCTCTGACAGGGTAATTTCAGTGGGAACTCCAGGGAGGTGATATGT	8689
Db	55576	GCACAAGAGAGATTCAACAAG-AGAGACCGTATTAAGTCAAAATAATTAAGCTCTTAAT	55634
Qy	8690	CAGTCACCGCTGTATATCTCATACACGGCTAATTAAGAACTTGTGTAGCAAGAAAC	8749
Db	55635	TCAATGATGAGATGTAAAGATTAATATGTTCTCTGCAATTAATTAATTAATAGT	55694
Qy	8750	ACTAGGTAATTAATCAACAATAATTTGTTAGACATCTGATTAAGAGTGGCATTTGCTTA	8809
Db	55695	AATCATTAATTAATAGCATAGCTTAGAAATCTCCAAAGTGTGGGAAAT-TAATGGCACAC	55753
Qy	8810	GGCATCAGATATACAGTAGTCAACATGGCAGACAAAGATGCTGCTGACAGGCTTGCTA	8869
Db	55754	TTCTTTATTTATCATGCTGTACAGGTAAATACACAAAGATTAAGAAACATTTGAAGTGA	55813
Qy	8870	AAATGAGAGAGCATTAAGAAAGAAAGAAAGAAAGAAATTAATTTAGGTATATTA	8929
Db	55814	ATGATTAATTAATAACGTAATGAATTAATGAAATACCTTAAGAGCTTTGGGAAAAGTT	55873
Qy	8930	AGGGTGTAAAGAAATTAAGACAGATAGTGGGATAGAGTGAGGAGATAGGGCTGTC	8989
Db	55874	TTCATATTTAAAGCTTACTGTAAABAAAAGAAAGATTTTAAATCAATCATATTTAA	55933
Qy	8990	TTCTGAAGAAATGATTTTGAAGCTGAGACTTCAGTGATGAGAGAAATTAACACACGAT	9049
Db	55934	CCAGAGATTGACAGACTTTTGTGTAAAGCTCAAGGTACAAATATTTGAGGCTTTGTA	55993
Qy	9050	GTCGTGAGAGAAAAGCATTTTAAAGGAGGTGAGCAGACATATCTTCAAGGAATCAAGAG	9109
Db	55994	GTCATGTAACTCTCATATTT--CCCTCTGGCCCAACATTTAAATAATGTAAATATCAG	56051
Qy	9110	GAAGCTGTGAGGCTGGAACACAGAGAAAGAGAGGTGGGTGACTTGAAGAGGACGGGA	9169
Db	56052	GCCAGCGTATGAGCTCATGCTGTATATCCACGACTTTGGGATGCTGAGCGAGGTGAT	56111
Qy	9170	CGGCAGTGGCCAGGTTAACTAGACCTGTGA-----AGGTTTTCACATTAAGAGGA	9222
Db	56112	CAGTTTGAATCAGAAAGTTCAAGACAGCCTGACCAAGATGTGTGCAACCCCATCTCTGCTA	56171
Qy	9223	GTCATCAGAAAGTCTTGAAGCAGGGCTGTGATTAATTTTAATCACTATTT-----T	9271
Db	56172	AAATAATCAAAAGTAGCCGAGATGCTGAGCATGCGCTCTAATCCACAGCTACTGGAGGCT	56231
Qy	9272	TATTAAGATCATCTGTCACTTTTTCAGAAACATTAAGTTATTAAGATCAAGCATTTAGC	9331
Db	56232	GAGCGAGAGAGTTGCTTGAACCTTGGAAAGCAGAGGTTGAATGTAGCCAAAGTACACCA	56291

QY	9332	AAGCAATCCAGC-----TACCAATCCGTGACGTTGTCMAAATTAGAGGTG	9376
Db	56292	CTGCATCTCCAGCCTGGGTGACAGAGTAGGCACTGTCTCAAAAACCAACAAAGAAACAA	56353
QY	9377	ATGACCCGCTTGAGCTAGAGATGATAGCAGAGAGGTGTGAGAAATCAACATGATATATTT	9436
Db	56352	CAAAAAACCAATTTTATAGGTAGGAAGCAATATATTAAGAACATAGGGCATCTTAGTATTT	56411
QY	9437	TGAGGTAGAGCTGACAGCATTTAACTAATAGCTAAGTAGGCCGGGTGTGTGGCTTAGG	9496
Db	56412	TGGG--AACCGTATCTGTGTTCATCTTAAAGATTCTAGCCGGGGCGGTGCTCACCC	56465
QY	9497	CTGTAACTCTAGCACTTTGGAGGCCAGGCGAGGTGATCACTGAGGTCAAGAGTTGG	9556
Db	56470	GCTGTGATCCAGACATTTGGAGGGCGAGGACAGGGCGATCACTGAGGTCTGGACTTTGG	56523
QY	9557	AGACCAAGCTTGACCAACATAGGTGAACCTGTCTCTAATAAAATTC--AAATAATGCTGG	9614
Db	56530	AGACCAAGCTTGACCAACATAGAGAAACCTGTCTCTAATAAAATTCAGCA	56588
QY	9615	GGAATGTGTGGCACTGCTGTAAATCTCAGGCTTACTGGGAGGCTGAGGACAGAGATCCG	9674
Db	56590	GGGGTGTGGCGCATGCTTATATCCAG--CTACTGGAGGCTGAGGACAGGAAACCGC	56648
QY	9675	TTGAACCTGGAGGTGTAATTTTGACGTAGACCGAGATTGACCACTTGACCTCAGCCTGG	9734
Db	56649	TTGAACCCAGGTGGCGAGGTGACGTGACGCAAGATTGACCACTTGACCTCAGCCTGG	56708
QY	9735	GGAACAGAGTGAACCTCCGCTCTTAATAAATGA	9770
Db	56709	GCAACAGAGTGAACCTCCATCTCAAAAAA	56744
RESULT 14			
HSBA12M9/c			
LOCUS	HSBA12M9	176287 bp	DNA linear PRI 04-MAR-2003
DEFINITION	Human DNA sequence from clone Rpl1-12M9 on chromosome 22, complete sequence.		
ACCESSION	AL080243		
VERSION	AL080243.21	GI:58706222	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 176287)		
TITLE	Pelam.S.		
JOURNAL	Direct Submission		
COMMENT	Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Sep 12, 1999 this sequence version replaced gi:5777540. ----- Genome Center		
	Center: Wellcome Trust Sanger Institute		
	Center code: SC		
	Web Site: http://www.sanger.ac.uk		
	Contact: humquerry@sanger.ac.uk		

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30): an attempt was made to resolve all sequencing problems, such as compressions and repeats, all regions were covered by at least one plasmid subclone or more than one M13 subclone, and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22> RP11-1293 is from the library RP11-11.1 constructed by the group of Plierer de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6. (*GenBank*)

FEATURES	SOURCE
Location/Qualifiers	1. 176287
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="RZPD:RPICB753M0912"	
/db_xref="taxon:9606"	
/chromosome="22"	
/clone="RP11-12W9"	
/clone_1b="RPC1-11.1"	
BASE COUNT	45020 a 40451 c 40308 g 50508 t
ORIGIN	

Query Match	4.44;	Score 439.6;	DB 9;	Length 176287;
Best Local Similarity	61.84;	Pred. No. 1.9e-105;		
Matches 909; Conservative	0;	Mismatches 484;	Indels 79;	Gaps 10;

QY	7182	TATTTAAAAATATCCAGGCGCGGCGTGGAGTCAAGCCCTGTATCTCTAGCACTTTGGGAG	7241
Db	130093	TTTTTAAACCCCAACAGCCAGGGGGTGGCTCAGCCTGTATATCCCAACATTTGGGAG	130034
QY	7242	GCTAGGCAAGGTGATTTGCTTGAAGCGGCTTGAAGCTTGAAGTTTGAACCAAGCTGGGCA	7301
Db	130033	GCCAGGGGGGTGAT-----CATAGGTCAGAGATTCAAGACCAAGCTGGGCA	129985
QY	7302	ACATGGCAAAACCTCATCTCTTCAAAAAATTTAAATTTAGCGGGGTGGTGAAGTCGTT	7361
Db	129984	ACATAGTGAACCCCATCTCTTCACTTAAATATCAAAAATTTAGCGGGGTGGTGGCATGTG	129925
QY	7362	CCTATAGTCCCATCTACTTCTAGAGG-----TGAGCCCGGAAGTGTG	7403
Db	129924	CCTATAGTCCCATCTACTCAGAGGCTGAAGAGAGAAATGCTTGAACCGGGAGGCGG	129885
QY	7404	AGGCTTCAGTAGCCGTGATCTGTGCTACTGCACTTCAGCTGGGTGACAGAGTGAACCA	7463
Db	129864	AGGCTGCAGTGATGGAGACTGCACACCTGCACCTCAGCTGGATGACAGAGTGAAGCTC	129805
QY	7464	TGTCCTCAAAAAACAAAAACAAAAACAAAAACAAACAAACAAACAAACAAACCA	7523
Db	129804	CATCTCAAAAAAGAAACAAACAAACAAACAAACAAACAAACAAACAAATTCATATTGAATTTT	129745
QY	7524	TATATATATATATATA-----CCTAGCTGAGGTGAGAAATGCACTATTTTGGTAA	7573
Db	129744	TTTTTTTGAATGCAATTTTCACTCTGTGGGCCAGGTGCAATCTCAGCTCATGGCAAC	129685
QY	7574	ATCAACCAATGACCCAGACTACAGATGGGAGCAATCCCTCCCTCTCACTGTGA--AATT	7631
Db	129684	TCCGCTCCCGGGTTCAGAGATTCCTCTGCTCAGCTCCCGATGCTGGGACTACAG	129625
QY	7632	TTTCTTTCTGTGACTCAAGTTTGTGTGTGTGTGTGTGTGTGTGATGGAGTCTCA	7691
Db	129634	GCGCTGCAACCAAGCCCGGCTTTTTTTTTTTTTTTTTTTTTTTCTGAGATGGAATCTCG	129585
QY	7692	CTCTGTCAACCCAGGCTGAGTGCATGGCGCAATCTTGTTCATCTGCAACCTCTGCTCC	7751
Db	129564	CTCTGTCAACCTAGAGCTGAGTGCAGTGGCAAGATCTCGGCTCACTGTAGCTCCGCTCC	129505
QY	7752	TGGGTTTCAAGGGAATCTCTCTGCTCAAGCTCCCGATATGCTGGGACTTCAAGCGGCAATCC	7811
Db	129504	TGGGTTTCAAGGCAATCTCTCTCTCAAGCTCTCTGTAGTACTGGAGACTTCAAGCGGCGCGC	129445

```

QY 7812 ACCATGCTGGCTAATTTTGTATTTTGTGGGTACAATGTAATTAATTAATTA 7871
DB 129444 ACTACGCCCGCTAATTTTCTGTATTTTATAGTAGACGGGGT-----TAAT 129397
QY 7872 ATTTTGTATTTTATAGTAGATAGGTTTCAACATGTTGGCCAGGCTGCTGAAATC 7931
DB 129396 TTTTGTATTTTATAGTAGACGGGGTTCACCGGTGTACCCAGGATGTTTCAATCTC 129337
QY 7932 CTGACCTCAGATGATCCGCTGCTCGGCTCCCAAGTGTAGATTAACAGCATAGC 7991
DB 129336 CTGACCTC--GTATCTGCGCCGCTTGTGCTCCCAAGTGTAGATTAACAGCATAGC 129279
QY 7992 AACGAGCCCTGGCCCTCAATAGTTTATCTATCTTCTTCTTCAACTTTGGC 8051
DB 129278 CACCGTGGCCCGGCGATCCAG-----CTAATTTTGTATTTTATAGTAGATGGGG 129227
QY 8052 TTGCAAGTGAACATCATGTTCTCTCACTTTCTCACTTCAATGATCTTTCAATCTCA 8111
DB 129226 TTTTCAACATGTTGGCCAGGCTGCTTGAATTTCTGACCTCAAGTAACTCAACCTGCTCA 129167
QY 8112 GTTCCAACTGATACCTCCCTCACTGCTCTTTTCTTCTAGATTTTCAAGAGAGGAA 8171
DB 129166 G-----CTTCCAAAGTGTGAGATTAACAGTGTAGGCCACTGCCAGCCT 129119
QY 8172 TCTGAATGCCCCAGTCCATATTTTCAACACACACATTAAGTGTGATTTGCCAGCC 8231
DB 129118 TAAGTCTTGAATTTTAAAGAGTGAAGTGTCTTAAATGATCAAGAAATGATCACTGGGGCC 129059
QY 8232 TAATGATGCTACATTAATAGGTTGGAACTCATCTTACTTCACTTGCACAAGAGAC 8291
DB 129058 A---TGTTTGAATAATTCATATGACATTAATTCACCTGCTCTTTTCAATGTATTAATA 129002
QY 8292 ATAGCTCTGCTGCTCAAAATAGGGCCCTGGGCCAGGTTGATGCTCATGCTATATAC 8351
DB 129001 GTGACACTTCTTGTGAAAATTAATCAATGAGCCGGAATGATGCTCAACCTGTAC 128942
QY 8352 CCAACTGTGGAGCGGAGGGGGGAGATCACTTGAATCAAGAGTTTGAACAGCC 8411
DB 128941 CCAAGCACTTGGAGGCGGAGGCGGAGATTCGCTGAGTCAAGAGTTGAGAGTCAAGCC 128882
QY 8412 TGGGCAACATGTTGAATCTCATCTCTCACTTAAATTAATTAATTAATTAATTAATTA 8471
DB 128881 TGGCCAACTGAGCAAAACCTGTGTCTACTTAAATAATC-AAAAATTAAGCTGGGTGTG 128823
QY 8472 GCATGACACAGTATGCTCCAGCTGTTCAAGAGCTGAGTGGAGATGTTGCTGAGTGTG 8531
DB 128822 GCGGCGGCTGTATATCCAGTACTTGGAGGCTGAGGCAAGAAATCACTTGAACGGG 128763
QY 8532 GAGGCAAGATGTCAGTGAACGCTGACTGTGCTTGTCAATCAAGCTGAGTGAACAGATT 8591
DB 128762 GAGGCAAGAGTGTCACTGAGCCAAAGATTACACCACTGCACTCGAGCTGGGCAACAGAGC 128703
QY 8592 GAGACCTGCTCAAAAAAATAATTAATAA 8623
DB 128702 AAGACTCTGTATGAAAAAATAAATAA 128671

```

```

RESULT 15
MMHMPX1 8852 bp DNA linear ROD 29-JUL-2002
LOCUS M.musculus gene for hemopexin (exons 1-7).
DEFINITION X56829
ACCESSION X56829.1 GI:22022645
VERSION haem binding; hemopexin; plasma glycoprotein.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1. (bases 1 to 925; 3296 to 4758)
AUTHORS Alam, J. and Smith, A.
TITLE Nucleotide sequence of the mouse hemopexin gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 8852)

```

```

AUTHORS Alam, J.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1990) Alam, J., Alton Ochsmere Medical Foundation,
Department of Molecular Genetics, 1516 Jefferson Highway, New
Orleans, LA 70121, USA
FEATURES
Source
1. 8852
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
join(<1..110,244..302,385..456,740..858,3355..3508,
4042..4254,4511..4639)
/label="mRNA_hmpex"
<1..110
/number=1
join(28..110,244..302,385..456,740..858,3355..3508,
4042..4254,4511..4639)
/codon_start=1
/product="Hemopexin"
/protein_id="CA40160.1"
/db_xref="GI:22022646"
/translation="MARTVALNIVLGLCSLAVASPLPIAHGVAVENGTKDS
DVEPHCLDWSFDAATMDNGTMLPFKGEFVRGHSSTBELSARWKNITVDAFR
GDSVFLIKEDKVMVVPKESKNGYKLFQBERPGIPIPPDAVCEHRESCSGVLF
FOGNRMFMDPATRTOXERSWPAVGNCTALRWLERYCYFOGNKFLRPVPTGEVPR
YPLDANDYFVSCGRGHRPRNGTAMGNSHTHPHSCRSPDPLSLSDHRTVAFT
"
111..243
/number=1
244..302
/number=2
303..384
/number=2
385..456
/number=3
457..739
/number=3
740..858
/number=4
859..3354
/number=4
3355..3508
/number=5
3509..4041
/number=5
4042..4254
/number=5
4255..4510
/number=6
4511..4639
/number=7
BASE COUNT 2362 a 1956 c 2106 g 2428 t
ORIGIN
Query Match 4.4%; Score 439.4; DB 10; Length 8852;
Best Local Similarity 54.8%; Pred. No. 1.8e-105;
Matches 2102; Conservative 0; Mismatches 1261; Indels 476; Gaps 43;
QY 3360 TTCTTGGGATTCATCTAGCCTTTCATATACCTTCCCTGAGTATCTCAACATCAC 3419
DB 2313 TCTTGGGGAACACATCAGACATTTTATGACCTCCCTGGGTGATTTAGAGCTCTGC 2372
QY 3420 CTTGCTGAGAGATATCTAGCTGTGACACATGATCTAGATCTTCTGAGCTCC 3479
DB 2373 CATG---CTTGAAGGCTCTGTGCTCTAGATACAGATCTGATCACTCCCTGAGCTTT 2429
QY 3480 AAGACATCTCTTCAATGTATGTTCTA-----CAGGACCTAAATTC 3523
DB 2430 GGAATCACTTTCACTATTTTGTGAAATCTGAGCTTTCCCAATACCTCAAGTCC 2489
QY 3524 AGCATCCCCAACTAGCTTGCATCTTCTTAAACCAACCTTCTCTGTGTTTC 3583

```


[illegible]

Db	5486	AAAT-----AGAAAATACCTTTTAACTTGAATGACTAAGCAAAATAATGATATTGTCTG	5538
Oy	6752	ACTATCAAATTCTGTGAAAAATACATCACCATGGAGAGCAAAATCTCCACAGCAGAGAT	6811
Db	5539	ATTTAACAAGTGCTGTGTATAAAGTGTGCACATCACTGGAATGTAGCT-----CAGCATGAA	5593
Oy	6812	TGCACACATATAATAGAACAATACGAGCTTAAGTGAACACACACCTGTATAGTAAATTCAA	6871
Db	5554	TGCACACCCTTATATGAAMAACATACGCTGCTGCANTTAAACATTAAGA--TGCTGTGCATATGCG	5651
Oy	6872	CATTAACTGAGAACATACGCCCATAGTAGTAAGAACATATGATATCA-----AGAGAACACA	6936
Db	5652	ATCTAGAGCT--GATTATATGTCAATATTTAAACATGTAAAGGTGAGAGATTATGAGTCTA	5709
Oy	6927	CAGCCATGCTGGAGGCCATTGGAGAGACACACAGACAAATGTGAATGCGAAMAAAGAGAGA	6986
Db	5710	TAGTAAGTGTGTAGATCACTGGGGAAACCAACAGAACAAAGTAAATGTATGTGTGTGT	5769
Oy	6987	GAGAGTAGTAGAGAGATTGTGAAAAACAGGGCCACAGCAAAACACACAGAAATATAGAGAG	7045
Db	5770	GTGTGTGTGTGTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	5828

Search completed: December 15, 2003, 23:41:40
Job time : 23561.6 secs

```

Db      5486 AAAT-----AGAAATACCTTTTAACTTCTTAATGACTAAGCAAAATATGATATGTTGTG 5538
Qy      6752 ACTATCAAGTTCTGTGAAAAATATCATCACCATGGAGAGCAAAATCTCCACAGCAGGAT 6811
Db      5539 ATTAACAAGTCTAGTGAAAAAGTATGCAACTACATGGAAATGTAGCT-----CAGCATGAA 5593
Qy      6812 TGCACTATATATAGAACTACAGCTTAAGTGAACACACACCTGTAGTGAATAATACAA 6871
Db      5594 TGCAACACCTTAATGAAAAACATACAGCTGCATTAATAACATAGA--TGCTGTGACAAATGCAG 5651
Qy      6872 CATTAAACTGGAACATACGCCATAGTAAGAACAATAGTATCA-----AGAGAACACA 6926
Db      5652 ATCTGAGCT--GATTATATGTCAATTAATTAACAATGTAAGATCAGAGATTATGAGTCTA 5709
Qy      6927 CAGCCATGTGTGGAGCCCATTTGGAGAGACACACAGCAAAAGTGAATGCAGAAAAGAGAGA 6986
Db      5710 TAGTAAGTGTGTAGATCACTGGGGAGAACACAGACAAAGTAAGTATGTGTGTGTGT 5769
Qy      6987 GAGAGTGAAGTGAAGATTGTGAAAAACAGGGCCACAGAAACACACAGAAATAGAGAG 7045
Db      5770 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5828

Search completed: December 15, 2003, 23:41:40
Job time : 23561.6 secs

```


XX Olek A, Piepenbrock C, Berlin K;
PI WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -
PS Claim 1; SEQ ID NO 917; 32bp + Sequence Listing; German.
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anemia, cancer, acute myeloid
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 5234 BF; 1479 A; 40 C; 1361 G; 2354 T; 0 other;
Query Match 15.1%; Score 1509.4; DB 24; Length 5234;
Best Local Similarity 79.9%; Pred. No. 3.3e-279;
Matches 1777; Conservative 0; Mismatches 446; Indels 0; Gaps 0;
QY 12 AGGAGGCCAGCAAACTCTGATGATTCAGACAGGGTCTGACAGCTGAAGCAATTTG 71
DB 3012 AGGAGGTTTATGATTAATTTGATGATTTAGATGAGGTTTGAATGATTAAGATTAAGTTG 3071
QY 72 TTGAGAAATTCCTGATGAGATCATGGGCTCTCAGAGGAGATATTAAGTTTCA 131
DB 3072 TTGAGAAATTTTGTATGAGAGATTAAGGGTCTTTAGAGGAGATTAAGTTTCA 3131
QY 132 AGGCTGAGAGGAAAGAAAGTGAAGGAGGAGCTTGAATAGTGCTCCATTTCCCA 191
DB 3132 AGGTTGAGAGGAAAGAAAGTGAAGGAGGAGCTTGAATAGTGCTTATTTATTTAA 3191
QY 192 CACCCAGAAAGAAAGCAATGCTCCATGAGGAGAGATGATGATTAAGTTTCA 251
DB 3192 TATTTAAAGAAAGAAAGTGAAGGAGGAGCTTGAATAGTGCTTATTTATTTAA 3251
QY 252 GCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 311
DB 3252 GTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3311
QY 312 TATGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371
DB 3312 TATGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3371
QY 372 GCGTATGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
DB 3372 GTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3431
QY 432 GTTGAAGAAATTCGACAGATGAGAAAGGAGCTGAGATTTTGTGAAGATTAATAT 491
DB 432 GTTGAAGAAATTCGACAGATGAGAAAGGAGCTGAGATTTTGTGAAGATTAATAT 491
QY 492 GGGATCTGACAGGCTGAGAGCATCTGTAGAGGATCAAGGCTTCAAGCTTGGCTAG 551
DB 492 GGGATCTGACAGGCTGAGAGCATCTGTAGAGGATCAAGGCTTCAAGCTTGGCTAG 551
QY 552 GAGAGAGGGGCTCTGAACTTCACTGCGCCATAGCTGAGCTGCGCCATATTTCTTC 611
DB 552 GAGAGAGGGGCTCTGAACTTCACTGCGCCATAGCTGAGCTGCGCCATATTTCTTC 611
QY 612 TGACTCTAGGCAATCTCAGACAGAAATGAGGACCTTGGAGTGGGCCAGGAAGT 671
DB 612 TGACTCTAGGCAATCTCAGACAGAAATGAGGACCTTGGAGTGGGCCAGGAAGT 671
QY 672 ACTGAGATAGCAGTGAATCCAGAGAGATGATGATGATGATGATGATGATGATGATGAT 731
DB 672 ACTGAGATAGCAGTGAATCCAGAGAGATGATGATGATGATGATGATGATGATGATGAT 731
QY 732 AGAAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
DB 732 AGAAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
QY 792 CATACCTACCCCAATTCCTATATTTCTCTCCTCCTAGAGGAGGAGAAATTTGACAAA 851
DB 792 CATACCTACCCCAATTCCTATATTTCTCTCCTCCTAGAGGAGGAGAAATTTGACAAA 851
QY 852 TTGGCTGACACCTGACCACTACTGATGATGATGATGATGATGATGATGATGATGATGAT 911
DB 852 TTGGCTGACACCTGACCACTACTGATGATGATGATGATGATGATGATGATGATGATGAT 911
QY 912 TTGAACAAATTTCTGGAACCACTCTTATGAGGAGGAGCTTTTGAAGTCAAGGAT 971
DB 912 TTGAACAAATTTCTGGAACCACTCTTATGAGGAGGAGCTTTTGAAGTCAAGGAT 971
QY 972 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
DB 972 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
QY 1032 GTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
DB 1032 GTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
QY 4032 GTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4091
DB 4032 GTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4091
QY 1092 ATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
DB 1092 ATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
QY 4092 ATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4151
DB 4092 ATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4151
QY 1152 AGCTAGAAATTTGTAATTTCCAGTATGATGATGATGATGATGATGATGATGATGATGAT 1211
DB 1152 AGCTAGAAATTTGTAATTTCCAGTATGATGATGATGATGATGATGATGATGATGATGAT 1211
QY 4152 AGTTAGAAATTTGTAATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 4211
DB 4152 AGTTAGAAATTTGTAATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 4211
QY 1212 GCTTTTCCCCAGGTGAACTACAAAGAAAGCTAATTAAGTATGATGATGATGATGATGAT 1271
DB 1212 GCTTTTCCCCAGGTGAACTACAAAGAAAGCTAATTAAGTATGATGATGATGATGATGAT 1271
QY 4212 GTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4271
DB 4212 GTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4271
QY 1272 TCTGAGAGAGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331
DB 1272 TCTGAGAGAGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331
QY 4272 TTTGAGAGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4331
DB 4272 TTTGAGAGAGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4331
QY 1332 CTTTGTGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1391
DB 1332 CTTTGTGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1391
QY 4332 TTTTGTGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4391
DB 4332 TTTTGTGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4391
QY 1392 TTTGACACAGTGAATTTATTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
DB 1392 TTTGACACAGTGAATTTATTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
QY 4392 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4451
DB 4392 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4451
QY 1452 GCGTTGAGGATGCGCCAGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1511
DB 1452 GCGTTGAGGATGCGCCAGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1511
QY 4452 GCGTTGAGGATGCGCCAGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4511
DB 4452 GCGTTGAGGATGCGCCAGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4511
QY 1512 TTTCTAGATGATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1571
DB 1512 TTTCTAGATGATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1571
QY 4512 TTTTCTAGATGATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4571
DB 4512 TTTTCTAGATGATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4571
QY 1572 GCTGCTGAGAGGCTTGAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1631
DB 1572 GCTGCTGAGAGGCTTGAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1631
QY 4572 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4631
DB 4572 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4631
QY 1632 CTCCCACTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691
DB 1632 CTCCCACTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691
QY 4632 TTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4691
DB 4632 TTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4691
QY 1692 TGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1751
DB 1692 TGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1751
QY 4692 TGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4751
DB 4692 TGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4751
QY 1752 TACCCAGTACTGCTTGAATTTGCTGAAGGAGGTTGGGAGGAGCTGCGGAGAAAG 1811
DB 1752 TACCCAGTACTGCTTGAATTTGCTGAAGGAGGTTGGGAGGAGCTGCGGAGAAAG 1811
QY 4752 TAGTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4811
DB 4752 TAGTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4811

QY 1812 GAGCTTGGATTGAGATTCTGTCAGACCCCTGACCTTATTTGCACTGATGATCAGCC 1871
 DB 4812 GAGCTTGGATTGAGATTCTGTCAGACCCCTGACCTTATTTGCACTGATGATGATGAT 4871
 QY 1872 AATATTGGCTTAATCTCGGAGACAGACATTCCTCAGTAGAGTTGGAGTGGGGTGGT 1931
 DB 4872 AATATTGGCTTAATCTCGGAGAGATGATATTTTATTTAGTAGAGTTGGAGTGGGGTGGT 4931
 QY 1932 CTGCTGCAACTATATAGGAGGATCACTGCTGACCCGAGCTGCTGCTGCTGCTG 1991
 DB 4932 TTTGTTGTTATTTATATATAGGAGGATTTATTTGTTATTTAGATTTGTTGTTGTTT 4991
 QY 1992 CAGCTCAGCATGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGG 2051
 DB 4992 TAGTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5051
 QY 2052 TGGTCTCTGCGCATTCGACCCCTCTTCTCTCTGAGTAAAGTGGACTAGAAAGCGAAG 2111
 DB 5052 TGGTCTCTGCGCATTCGACCCCTCTTCTCTCTGAGTAAAGTGGACTAGAAAGCGAAG 5111
 QY 2112 GATTGATCTGCGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGG 2171
 DB 5112 GATTGATCTGCGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGG 5171
 QY 2172 TCAGGGGCTATGAGGAAAGGAGTGGTCCCAATGATCAAGATATCTATTTGTTCCCT 2231
 DB 5172 TTAGGGGCTATGAGGAAAGGAGTGGTCCCAATGATCAAGATATCTATTTGTTCCCT 5231
 QY 2232 AGG 2234
 DB 5232 AGG 5234

RESULT 2
 ABL32945/c
 ID ABL32945 standard; DNA; 5234 BP.

XX ABL32945;

DT 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 918.

XX Human: immune system disease; cytosine methylation; antiaesthetic;
 KW antiarteriosclerotic; antiaesthetic; cytosine methylation; antiaesthetic;
 KW neuroprotective; anti-HIV; anticonvulsant; opthalmological;
 KW antineuritic; antineuritic; antidiabetic; antiparasitic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; dr.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JUN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Plegenbrock C, Berlin K;

XX MPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX

PS Claim 1; SEQ ID NO 918; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.

XX Sequence 5234 BP; 1303 A; 40 C; 1091 G; 2800 T; 0 other;

Query Match 12.8%; Score 1281.6; DB 24; Length 5234;

Best Local Similarity 73.4%; Pred. No. 1.3e-235; Matches 1638; Conservative 0; Mismatches 594; Indels 0; Gaps 0;

QY 1 TCCCTCTCCCGAGGCGCCAGCAAAATCTGAGATTGAGACAGAGGTTCTGACAGCTG 60
 DB 2234 TCCCTCTCCCGAGGCGCCAGCAAAATCTGAGATTGAGACAGAGGTTCTGACAGCTG 2175
 QY 61 AAGACAAGTTGTTGAGGAAATCTGATGAGGATCATGGGCTCTCAGAGGAGAGATA 120
 DB 2174 AAAACAATTTATTTAAAAATTTCTTAAATAATCATTAATTAATCAAAAAATAA 2115
 QY 121 TAAGTTTGAAGGCTGAGAGGAGAAAGAGTGAAGGAGTCTTAATAGTGGCTC 180
 DB 2114 TAAATTTCAAAATCTTAAAAATTTAAAAATTTAAAAATTTAAAAATTTAACTC 2055
 QY 181 CCATTGCCCAACCCAGAAAGAGACATGCTCCGATGAGGAGAGGATGATAGCA 240
 DB 2054 CCATTGCCCAACCCAGAAAGAGACATGCTCCGATGAGGAGAGGATGATAGCA 1995
 QY 241 CATTGCTGAGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 300
 DB 1994 CATTGCTGAGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1935
 QY 301 GCGCACTGACCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 DB 1934 ACCCACTGACCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1875
 QY 361 GAGGACATTGAGGCTGATGAGGAGGCTGATGAGGAGGCTGATGAGGAGGCTGATGAG 420
 DB 1874 AAAAAATTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1815
 QY 421 TGTGTTTCTGTTGAGAGGAACTTTCGAGAGTAAAGGAGTGAAGCTTTTGGTA 480
 DB 1814 TATATTTTCTTATTAATAAAATTTTCGCAAAATAAATAAATAAATAAATAAATAA 1755
 QY 481 AATTTATATATGAGCTGACAGGAGTCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAG 540
 DB 1754 AATTTATATATGAGCTGACAGGAGTCTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAG 1695
 QY 541 GCGTGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 DB 1694 ACCTTAATTAATAAAATTAATAAAATTAATAAAATTAATAAAATTAATAAAATTAATA 1635
 QY 601 TAATTTTCTGACTCACTAGGCAATCTCACAGAAATGAGGAGGAGGAGGAGGAGGAGGAG 660
 DB 1634 TAATTTTCTGACTCACTAGGCAATCTCACAGAAATGAGGAGGAGGAGGAGGAGGAGGAG 1575
 QY 661 GCGCAGAGAGTACTAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 DB 1574 ACCCAAAAATTAATAATAAAATTAATAAAATTAATAAAATTAATAAAATTAATAATA 1515
 QY 721 ACACTGAGAGAGAAACAGAGCTGTCAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
 DB 1514 ACACTGAGAGAGAAACAGAGCTGTCAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1455
 QY 781 ATGCTTACTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
 DB 1454 ATGCTTACTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1395

QY	841	ATGTCAGAAAATTGGCTGCACACATAGAACACACTACAGACTGTGAATGCAATTTTG	900
Db	1394	ATTATCAAAAAATTTAACACACACTACACACACTACACACTTAATCAATTTTA	13334
QY	901	CATTTTTTCATTCAACAAATATTTCTGGAAACACTCTATATGCGAGCACTATTTTG	960
Db	1334	CATTTTTTCATTCAACAAATATTTCTAABAAACACTCTATATATCAACACTATTTTA	12756
QY	961	GAGTCAGGGATATATATATGTTAAACAAGACGGCAAAACAAAGCAAGCACACACCA	10222
Db	1274	AAATCAAAAAATATATATATATATATTAACAAACAAACAAACAAACAAACAAACAA	12151
QY	1021	TCACCAATTAAGTAGACAGATGAAGAAATTTCAAGTTTATAGTAATTAATTAAC	10806
Db	1214	TCACCAATTAATATACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC	11552
QY	1081	AAGGCTGAATGCGCTAGATTAAGCGGTCAGAAAGGCTTCATTGAGAAAGTAGATTT	11448
Db	1154	AAAAATCTAATAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	10995
QY	1141	AAGCAGAGTCAGCTAGAAATTTGTGAATTCAGTTACAGTTCTATTTGTCTGGTT	12007
Db	1094	AAACAAAAATCAACTAAATATATTTTAATTAATTCCAATTAACAATCTATTTATTC	10351
QY	1201	GGTTAATTAAGCTTTTCCCCCAGAGTGAATACACAGAAAGACATTAATTAAGTAG	12606
Db	1034	AATTAATTAATTAATTTTCCCCCAAAATTAACAACTACCAAAAAAATCTAATTA	975
QY	1261	TGCTGCTCTCTGGAAGAGACACCTCTGTTCTGCTCATTACTGTCAACCTTC	13202
Db	974	TAAATATATCTCTTAATAAAAAAACAACCTCTGATTTCTACCTCATTAATCAACCTTC	915
QY	1331	ACTTCAGGCACTTTTGGCAAAAGCCCTTTGCCAGTCAGGAAAGCGGAGGCTGGCATG	13808
Db	914	ACTTCCAACACTTTTTCACAAACCCCTTACCAATCAAAAAAAGAAAACTAACAATA	855
QY	1381	GGGCTTGACATTTGACAAACGTGACATTTGTGCCCGACCTCACTAGCCCAAGGT	14410
Db	854	AAACTTAACATTTAACAACATTAATAACATTTATATCCCAACCTCAACCTCAACCT	795
QY	1441	AAAGCTGAAGGCTTGGGCATGCCCCAGAAAGGCCCTGATGAAGTTGGAAAAAGCTG	15007
Db	794	AAAACTAABAAAAACTTAACACATTAACCCCAAAAAAACCTTAATTAATTAATTAAT	735
QY	1501	TTCTCTGAGTATTTCTAAGTAAATTTCTGTGTGTGTGTACTTAAGTAGTAGAT	15606
Db	734	TTCTCTAATATTTCTAATTAATTAATTTATCTATATATATTAATTAATTAATTAAT	675
QY	1561	TGCTGTCTTAGCTGCCCTTAGAGCAGGGCTTGACACAGTACACAGCAATTAATGCCCT	16220
Db	674	TACTATCTCTAATCTAATCTTAATAAAACAAACCTTAACAAATTAACAAATTAATTAAT	615
QY	1621	CCTTTTCTCACTCCCACTTGTGGAGATTAATCAATCAACAAAGGTGATCTCACT	16808
Db	614	CCTTTTCTCACTCCCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	555
QY	1681	ACTCACTTCCCTGACTTAATGATGCTGAGCCCATTTGCAAGTGTGAGTCAACGTGGA	17404
Db	554	ACTCACTTCCCTGACTTAATTAATTAATCTTAACCCATTAACAAATTAATTAATTAAT	495
QY	1741	CGTAGCAGTGTAGCCCACTTAATCTGCTTGAATTTGCTGAAGGGGGTGGGGGGCACTG	18000
Db	494	CGTCACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	435
QY	1801	CCGGGAAAAAGAGCTTGATGATTCAGATTTCTGTCAACCCCTGACTTAATTTGCAAGTA	18666
Db	434	CCGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	375
QY	1861	TGTATACGCCAATATTTGGCTTAATGCTCTGGAGACAGACATTTCCCACTAGAGTTGAGG	19202
Db	374	TATATATCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	315
QY	1921	TGGGGGTGTGTCTGCTGCCAACTCTATATATAGGAAGTTCACTGTCAACCCAGAGCTGTC	19808

Db	314	TAAATAATATCTACTACCACTCTATATATAAAATTTCAATATATCAACCCAAAATCTATCC	255
Qy	1991	TGTGGCTCTGCACTCAGCATGGCTAGGGTACTGGAGCAACCGTTGCACTGGGGTTGT	2040
Db	254	TATAAACCCTCAACTCAACATCAATCTAAAAATACTAAAAACACCCGGTACACTAAAAATAT	195
Qy	2041	GGAGCTATGCTGGCTCTGGSCATTTGCGCACCCCTCTTCCCGGTGAAGTAAGCTGGGAC	2100
Db	194	AAACCTATATCTAATCTCTTAACCAATACACCCCTCTTCCCGTAAATTAACCTAAAC	135
Qy	2101	TAGAAGCGAAGATTAAGTTCTGGGCTAGGGTAAGTAAGGCGCAATTTTTAGGCTCGGT	2160
Db	134	TAAACGAAAAAATTTAATTTCTAATTAATTAATTAATTAACCAATTTTTTAACCTGAT	75
Qy	2161	CAAAATTTGGGGTCAAGGGGCTATGGGAAAGGAGTGGTCCCAATGATCAAGATATCTATT	2220
Db	74	CAAAATTTAAATCAAAAACTATAAAAAAATAATGATCCCAATTAATCAAAATATCTATT	15
Qy	2221	TTGTTCTCCCTTA	2232
Db	14	TTATTTCTCCCTTA	3
RESULT 3			
AAK87529			
ID	AAK87529 standard; DNA; 13608 BP.		
AC	AAK87529;		
XX			
DT	07-NOV-2001 (first entry)		
XX			
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:42341.		
XX			
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
XX	cytostatic; gene therapy; vaccine; metastasis; ds.		
OS	Homo sapiens.		
XX			
PN	WO200157182-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01354.		
XX			
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189074.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		

PR	14-AUG-2000	2000US-0225758	PR	08-NOV-2000	2000US-0246610
PR	14-AUG-2000	2000US-0225759	PR	08-NOV-2000	2000US-0246611
PR	18-AUG-2000	2000US-0226279	PR	08-NOV-2000	2000US-0246613
PR	22-AUG-2000	2000US-0226681	PR	17-NOV-2000	2000US-0249207
PR	22-AUG-2000	2000US-0226686	PR	17-NOV-2000	2000US-0249208
PR	22-AUG-2000	2000US-0227182	PR	17-NOV-2000	2000US-0249209
PR	23-AUG-2000	2000US-0227009	PR	17-NOV-2000	2000US-0249210
PR	30-AUG-2000	2000US-0228924	PR	17-NOV-2000	2000US-0249211
PR	01-SEP-2000	2000US-0229287	PR	17-NOV-2000	2000US-0249212
PR	01-SEP-2000	2000US-0229343	PR	17-NOV-2000	2000US-0249213
PR	01-SEP-2000	2000US-0229345	PR	17-NOV-2000	2000US-0249214
PR	05-SEP-2000	2000US-0229509	PR	17-NOV-2000	2000US-0249215
PR	05-SEP-2000	2000US-0229513	PR	17-NOV-2000	2000US-0249216
PR	06-SEP-2000	2000US-0230437	PR	17-NOV-2000	2000US-0249217
PR	06-SEP-2000	2000US-0230438	PR	17-NOV-2000	2000US-0249218
PR	08-SEP-2000	2000US-0231242	PR	17-NOV-2000	2000US-0249244
PR	08-SEP-2000	2000US-0231243	PR	17-NOV-2000	2000US-0249245
PR	08-SEP-2000	2000US-0231244	PR	17-NOV-2000	2000US-0249264
PR	08-SEP-2000	2000US-0231413	PR	17-NOV-2000	2000US-0249265
PR	08-SEP-2000	2000US-0231414	PR	17-NOV-2000	2000US-0249297
PR	08-SEP-2000	2000US-0231414	PR	17-NOV-2000	2000US-0249299
PR	08-SEP-2000	2000US-0232081	PR	17-NOV-2000	2000US-0249300
PR	12-SEP-2000	2000US-0231968	PR	01-DEC-2000	2000US-0250160
PR	14-SEP-2000	2000US-0232397	PR	01-DEC-2000	2000US-0250391
PR	14-SEP-2000	2000US-0232398	PR	05-DEC-2000	2000US-0251030
PR	14-SEP-2000	2000US-0232399	PR	05-DEC-2000	2000US-0251988
PR	14-SEP-2000	2000US-0232400	PR	05-DEC-2000	2000US-0256719
PR	14-SEP-2000	2000US-0232401	PR	06-DEC-2000	2000US-0251479
PR	14-SEP-2000	2000US-0233063	PR	08-DEC-2000	2000US-0251856
PR	14-SEP-2000	2000US-0233064	PR	08-DEC-2000	2000US-0251869
PR	14-SEP-2000	2000US-0233065	PR	08-DEC-2000	2000US-0251899
PR	21-SEP-2000	2000US-0234273	PR	08-DEC-2000	2000US-0251990
PR	25-SEP-2000	2000US-0234997	PR	11-DEC-2000	2000US-0254097
PR	26-SEP-2000	2000US-0234998	PR	05-JAN-2001	2001US-0259678
PR	27-SEP-2000	2000US-0235464	XX	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000	2000US-0235834	XX	Rosen CA, Barash SC, Ruben SM,	
PR	29-SEP-2000	2000US-0236337	XX	WPI; 2001-483426/52.	
PR	29-SEP-2000	2000US-0236337	XX		
PR	29-SEP-2000	2000US-0236338	XX		
PR	29-SEP-2000	2000US-0236359	XX		
PR	29-SEP-2000	2000US-0236370	PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	02-OCT-2000	2000US-0236802	PT	useful for preventing, diagnosing and/or treating cancers and	
PR	02-OCT-2000	2000US-0237037	XX	metastasis -	
PR	02-OCT-2000	2000US-0237038	PS	Disclosure; SEQ ID NO 42341; 3071bp + Sequence Listing; English.	
PR	02-OCT-2000	2000US-0237039	XX		
PR	13-OCT-2000	2000US-0239935	XX	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
PR	13-OCT-2000	2000US-0239937	CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic	
PR	20-OCT-2000	2000US-0240960	CC	activity, and can be used in gene therapy and vaccine production. (I)	
PR	20-OCT-2000	2000US-0241221	CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
PR	20-OCT-2000	2000US-0241785	CC	treatment of diseases associated with inappropriate (I) expression. For	
PR	20-OCT-2000	2000US-0241786	CC	example, they may be used to treat disorders associated with decreased	
PR	20-OCT-2000	2000US-0241787	CC	expression by rectifying mutations or deletions in a patient's genome	
PR	20-OCT-2000	2000US-0241808	CC	that affect the activity of (I) by expressing inactive proteins or to	
PR	20-OCT-2000	2000US-0241809			

Db	6883	GTTCAACATTTAATTAATTAATTTGTAATTTTAATGTTTTGTTGTTGTTGTTGATG	6942
Oy	7664	GAGTCTCACTC-TGTCACCCAGGCTGAGTGCATATGGGCAATCTTGTTCACTGCAACC	7742
Db	6943	GAGTTTCGTTCTTGTTGGCCAGGCTAGAGAGCAATGGGTATCTTGCTCACTGCACACC	7002
Oy	7743	TCGTGCTCTGGGTTCAAGGCATCTCTGCTCTGAGCTCCCGTAATGACTGGACTACAG	7802
Db	7003	TCGTGCTCTGGATTCAAGTATTTCTGTCTCAAGCTCTCGACACTGGGAGTTAAG	7062
Oy	7803	GCGCATACCAACATGCTGGCTAATTTTGTATTTTTTGGGTTACATATGTAATATTY	7862
Db	7063	GTGCCCGGCATACACCAAGTAATTT-----	7089
Oy	7863	ATTAATTTAATTTTGTATTTTGTAGAGATAGGTTTCAACATGTTGGCAGGCTGCT	7922
Db	7090	-----TTTGGTATTTTGTAGTAGACAGGGTTTCACTTGTGACTGCTGGCTGG	7139
Oy	7923	CTGCACCTCTGACCTCAAGTGAATCCGCTGCTCTGCTGCCCTCCCAAGTCTAGATTACA	7982
Db	7140	CTTGAATCTCTGACCTCAAGTGAATCCGCTGCTGCCCTCCCAAGTCTGGCAATTACA	7199
Oy	7983	GGCATAGACCAACAACGCTGGGCCCCCTAAGGTTTATATCTATTTCTTGTCTTCTTAC	8042
Db	7200	GCGGTAGACCAACCGGCCCCGCTTAATTTTATGATTTTAAAGACGAGCTTGTGCTCAT	7259
Oy	8043	AACTTTGGCTTGCACGTGGAACATCATGTTCTTCCACTTTTCTCATTTCAATGATCTY	8102
Db	7260	CACCAAGGCTGAGTGCAGTGGTGCAATTAACAGCTCATTTGAGC-----CTGAACTC	7312
Oy	8103	TCAGTCTCAATTTCCACTGATTAATCTCTCCCTCAGT-TGCTCTTTTTCCTAGTAAGATTTC	8161
Db	7313	CTGGGCTTCAGCAATTTCTTGTCCCTCATCTCTCAAGTGGGCACTACAGTAATAGT	7372
Oy	8162	AGAGGGGAATCTGAATGGCCAGTCATATTTTCAGACCAACACACTTAAGATGCTG	8221
Db	7373	CACCATGCCAGCTAATTTTAAATTTTCTGTAGAGCGAGTCTCATATGTTGCCCA	7432
Oy	8222	ATTGCAAGCTATGTAATGGCTACATTAATGGGTTGGAACTCATCAT---TCACTTCAT	8278
Db	7433	GGGTGTCTCAACTCTGGGCTCAAGGAACTCTTCCACTCTGCTCCCAAGTGTG	7492
Oy	8279	TGCACAAAGCAGCATAGCTCTGTGTTCTCAAAATATAGGCCCTGTGGCCAGGTGTGTGCT	8338
Db	7493	GGATTACAGGTATTAAGCCATCAAGCCCAAGTCTCAAAACATTTAAACAGGAGAGAGCT	7552
Oy	8339	CATGCGTATATCCCAACATGTTGGAGGCGGAGGGGCGAGTCACTGAGTCAAGAG	8398
Db	7553	CGACCTGTATATCCAGCAATGTGGAGGCTGAGGCAAGCGGAGTCACTTGATTCAGAG	7612
Oy	8399	TTCTAGACCAAGCTGGGCAACATGGTGAATCTCATCTCATTAATAAATAAATAAATA	8458
Db	7613	TTTGAACCAAGCTGAGCAACATGTTGAACCTGTCTCTACAAAATAATC-AAAAATTA	7671
Oy	8459	GCCAGGTGTGTGGCATGCAACCAAGTATGCCAGCTGTTCAAGAGGCTGAGGTGGAGAT	8518
Db	7672	GCCGGGCGCTGTGGCATATGCTCTGCAAGTCCGGGCTACTTGGAGCGTAAAGCTGGAGGAT	7721
Oy	8519	TGCTCTGAGTGTGGAGGACAGATTTGCAATGTAACCGTGAATCTGTGCTCTGCAATCCAGC	8578
Db	7732	GGGTTGAGCCCGAGGAGGTGAAGA-GGCAAGCAGCTGAATTTGCAACACTGCACTCCAGCC	7790
Oy	8579	TGGGTGACAGATTGAGACCTGTCTCAAAAAACAAATAAATAAATAAATATATGCT	8638
Db	7791	TGGGTGACAGACCGCAAGCTGTCTTCAATATCAACACACGAAAA-----	7835
Oy	8639	TCTTGAGCAGGTAATTTCAATGGGAAACTTCCAGGGAGGTGATATGTCATGCAACG	8698
Db	7836	-----ACAGGAGGGTAATTAAGAACCAAACTAGATCACTTATGTGTCTTAAGTATGT	7887
Oy	8699	CTGTATACTCAATACCGGCTAATTAAGAACTTGTGTAGACAGCAAGAACATTAAGTAT	8758

Db	7888	CTTACCAACCCAGTGCAGCTCT-----GAGAAATTAAGAAGCACTGGCAA	7990
Qy	8759	TTACTCAACAAATATTTGTGTGAGCATCTGATAGAAAGTGGCAATGTGTCCTTAGGCACTGAG	8818
Db	7931	GAGATGAAAGCTCTAGTTATTGCTCTCAACAGTTACTCAAGAAATCTCAGAGAAATGTCAT	7990
Qy	8819	ATTACAGTATGCAACATGGCCAGACAAAGATGCTGCGCTGACAGGCTCTGTAAGTGAAG	8878
Db	7991	TCACTCTAGTTCGAATCTCATTTCAATTAATCATTTGATGTCATTAAGAACTTTTCTTT	8050
Qy	8879	AGCACAATTAAGAAAGAGAAAGAAAGAAAGAAATTAATTTTACGTAAATTAAGGTTGTA	8938
Db	8051	AAAACAAAACAGAAATGACTTTTACAGTCATTTCACTTTGTTAG---AATATATTTATTA	8107
Qy	8939	AAGAAATTAAGACAGATAGTGGATAGAGTGAAGAAATGAGGCTGTCTTCTGAAGA	8998
Db	8108	CTGCTCGGCACATAGATGGCTGTCTTAATTATGGGATTTTAAGAGATCCGTCTCTACA	8167
Qy	8999	AATATATTTTGAAGTGAACCTTCAGTGAATGAAAGAAATTAACCAACAGATGTCGTGAG	9058
Db	8168	AAAAATTTTAAATTAAGCTGGGTGTGTGCACATGCTTGTGTCCAGCTATTCAGAG	8227
Qy	9059	GAAAGCATTTTGAAGAGGCTGAGACACACATCTTCAAGAAATCAAGAAAGAAAGCTGCG	9118
Db	8228	GCTGAGGTGAAATAACATCTTGTAGC--CAAGAGTCAAGATTACATGATCTTGAATTG	8285
Qy	9119	TGAGGCTTGAACACA-----GAGAAAGACAGGTGGGTGACTTTGAAGAGGACAGGACGG	9172
Db	8286	TGCCACTGTATCCCAAGCCTGGCCAAAGAGCAAGACCCTGTCTCAAAAAAATAAAAAA	8345
Qy	9173	CAGTGGCCAGTTACCTAGACCTGTGTAAAGGTTTCAACATTAAGAGGATCATCAAA	9232
Db	8346	GGCCTGGCGCGGTGACTCAAGCTGTAA--TTCCAGGCTTTTGGAGGCGGAAGTGGGTG	8403
Qy	9233	AGCTTTGAGCAGGGCTGTGATATATTTCTAACTATTTTATTAAGATCACTCTGACTT	9292
Db	8404	AATACATCTGAGGTTTAAGATTTGAGACACAGCTGCGCAACATGCGCAAAACCCATCTATA	8463
Qy	9293	TTTGCAGAACCTAAGTTATTAAGATACAGACATGTAAAGCAAGAAATCAGCTAGCAATCC	9352
Db	8464	CTTAATAATTAAGAAATTTGGCCAGGCAATGATGTGTCAATGCTGTATATCCCAAGTACTAG	8523
Qy	9353	GTGCATTTTTCCAAATTAAGAGTGTATGACCGCTTGAATCTAGATGATGACACAGAGTG	9412
Db	8524	GAGGCTGAGGTACAAAGATCACTTGAACAGAGGCGGAGGTTGTCAGTGAAGCCAGATT	8583
Qy	9413	GTGAGGAATCAACATGATATATTTTGAAGGTAGAGCTGACAGCAATTAATTAATGCTAA-	9471
Db	8584	GCACCACTGCAGTCCAGCCTGGGCAACAGACAGCAAGACTCAATCTTTAATAAACAAACAAA	8643
Qy	9472	-----GATAGGCGCGGTGTGTGTGGCTTAAGCTGTATCTAGCACTTTTGGAGGCGCAA	9525
Db	8644	CAAAACAAGAGGCTGGGCAAGTGGCTCAAGGCTATATATCCAGCACTTTGGGAGGCGGA	8703
Qy	9526	GGCGAGTGTGATCACCTGAGGTCAAGAGTTGAGACACAGCTTGAACCAATGTGTGAACCT	9585
Db	8704	GGCAGAGTGTATCACTGAGGTCAAGAGTTCAAGACACAGCTGACCAATGAGGAAACAC	8763
Qy	9586	CGTCTTAATTAATAATCAAAATTAAGCTGGGAATGTGTGGCAATGCTCCTGTAAATCTAGCC	9645
Db	8764	TGTCTTAATTAATAATTAATAATTAAGCTGGGCGGTGTGGGCAATGCTCCTGTAAATCTAGC	8822
Qy	9646	TACTTGGAGGCTGAGCAGAGAAATGCTTTGAACCTGGAGGTAATGTTTGCAGTAGC	9705
Db	8823	TACTTGGAGGCTGAGCAGAGAAATGCTTTGAACCTGGAGGCAAGAGTTTGCAGTAGC	8882
Qy	9706	CGAGATTGCAACATTTGCACTCCAGCTGTGGGAAACAAGATGAATCTCCGTCTTAATAAA	9765
Db	8883	CGAGGTTGTGTGCACTGCACTCCAGCTGTGGGAAACAAGATTAATCTCCGTCTTAATAAAA	8942
Qy	9766	ATGATTAATGATGATATCAATCA	9790
Db	8943	AAAAAAAAAAAAAAAAAGAAAGGA	8967

RESULT 4
AAS32492
ID AAS32492 standard; DNA; 32191 BP.
XX AAS32492;
AC AAS32492;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human genomic DNA for novel endocrine antigen, SEQ ID No 446.
XX
KW Human; endocrine antigen; ds; cytostatic; antifertility; antidiabetic;
KW thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
KW antineoplastic; antibody; endocrine disorder; hormone imbalance;
KW reproductive disorder; endocrine cancer; pancreatic disorder;
KW diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
KW hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.
XX
OS Homo sapiens.
XX
PN MO200155319-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01335.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216547.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234423.
PR 21-SEP-2000; 2000US-0234474.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.

QY	8380	GATCACTGATCCAGGAGTTCTTGACACAGCCTGGGCAACATGGGAATCTCACTCTTA	8439
QY	8380	GATCACTGATCCAGGAGTTCTTGACACAGCCTGGGCAACATGGGAATCTCACTCTTA	8439
Db	5205	GATCACTGAGGTCAAGGCTTCGAGACCAAGCCTGGCCAACTGGTAACCCCACTCTTA	5148
QY	8440	CTAAAAATACAAAAATTAGCCAGGTGTGTGTGGCATGCAACGATAGTCCACGTCTTTCAG	8499
Db	5145	CTAAAAATAC -AAAAATTAGCTGGGTGTGTGTGTGGCGCCTGTATCCACGCTACTCGG	5087
QY	8500	GAGGCTAGAGTGGGAGATATGCTCGAGTGTGTGGAGGCAAGATTTGCAATTTGAACCTGACT	8558
Db	5086	GAGTCCGGGAGAGGAATTCCTTTGAACCCAGAGGTGAGGTTGCACTGAGCTGAGATC	5027
QY	8560	GTGCTCTGCAATCCAGCCTGGGTGAC -AGATTGAGACCTGTCTCAAAAAACAATTTAA	8618
Db	5026	ACGCATTTGCACTCCAGCCTGGGTGACAAAGCAAAAACTCAGTCTCAAAAAAAAAAAGA	4967
QY	8619	TAAAAATAAAAATATATGTTC	8640
Db	4966	AAAAAAAAAGATATATATTTCC	4945

RESULT 6
ABA20857/c
ID ABA20857 standard; DNA; 11234 BP.

AC ABA20857;

DT 23-JAN-2002. (first entry)

DE Human nervous system related polynucleotide SEQ ID NO 13188

KM Human, ootrophic; neuroprotective; cytosolic; dermatological; vincicide;
 KM immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KM antiparkinsonian; antischlicking; antianaemic; antirheumatic; cancer;
 KM antineuronic; hepatotropic; cerebroprotective; antiinflammatory;
 KM antiallergic; antidiabetic; antiviral; anticonvulsant; antitumoral;
 KM antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; nephrotropic; gene therapy; vaccine; ds

OS Homo sapiens

PN WO200159063-A2

PD 16-AUG-2001.

17-JAN-2001; 2001WO-US013334.

PR	1-JAN-2000	2000US-0179065
PR	04-FEB-2000	2000US-0179068
PR	24-FEB-2000	2000US-0184664
PR	02-MAR-2000	2000US-0186350
PR	16-MAR-2000	2000US-0189854
PR	17-MAR-2000	2000US-0190076
PR	18-APR-2000	2000US-0196123
PR	19-MAY-2000	2000US-0205515
PR	07-JUN-2000	2000US-0209467
PR	28-JUN-2000	2000US-0214886
PR	30-JUN-2000	2000US-0215135
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216800
PR	11-JUL-2000	2000US-0214877
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220393
PR	26-JUL-2000	2000US-0230964
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270

PR	14-ANG-2000	2000US-02255447
PR	14-ANG-2000	2000US-02255757
PR	14-ANG-2000	2000US-02257558
PR	14-ANG-2000	2000US-02257559
PR	18-ANG-2000	2000US-02262719
PR	22-ANG-2000	2000US-02266681
PR	22-ANG-2000	2000US-02266688
PR	22-ANG-2000	2000US-02277182
PR	23-ANG-2000	2000US-02277009
PR	30-ANG-2000	2000US-02288924
PR	01-SEP-2000	2000US-02292887
PR	01-SEP-2000	2000US-02293943
PR	01-SEP-2000	2000US-02293944
PR	01-SEP-2000	2000US-02293945
PR	03-SEP-2000	2000US-02295009
PR	05-SEP-2000	2000US-02295133
PR	05-SEP-2000	2000US-02295437
PR	06-SEP-2000	2000US-02304348
PR	08-SEP-2000	2000US-02312462
PR	08-SEP-2000	2000US-02312463
PR	08-SEP-2000	2000US-02312464
PR	08-SEP-2000	2000US-02314133
PR	08-SEP-2000	2000US-02314141
PR	08-SEP-2000	2000US-02314142
PR	08-SEP-2000	2000US-02314060
PR	08-SEP-2000	2000US-02320811
PR	12-SEP-2000	2000US-02321968
PR	14-SEP-2000	2000US-02323397
PR	14-SEP-2000	2000US-02323398
PR	14-SEP-2000	2000US-02323399
PR	14-SEP-2000	2000US-02324000
PR	14-SEP-2000	2000US-02324001
PR	14-SEP-2000	2000US-02324002
PR	14-SEP-2000	2000US-02324003
PR	14-SEP-2000	2000US-02324004
PR	14-SEP-2000	2000US-02324005
PR	14-SEP-2000	2000US-02324006
PR	14-SEP-2000	2000US-02324007
PR	14-SEP-2000	2000US-02324008
PR	14-SEP-2000	2000US-02324009
PR	14-SEP-2000	2000US-02324010
PR	21-SEP-2000	2000US-02342723
PR	21-SEP-2000	2000US-02342724
PR	23-SEP-2000	2000US-02342997
PR	23-SEP-2000	2000US-02342998
PR	26-SEP-2000	2000US-02354884
PR	27-SEP-2000	2000US-02358934
PR	27-SEP-2000	2000US-02358935
PR	27-SEP-2000	2000US-02358936
PR	29-SEP-2000	2000US-02363569
PR	29-SEP-2000	2000US-02363570
PR	02-OCT-2000	2000US-02368082
PR	02-OCT-2000	2000US-02370337
PR	02-OCT-2000	2000US-02370338
PR	02-OCT-2000	2000US-02370339
PR	02-OCT-2000	2000US-02370340
PR	02-OCT-2000	2000US-02370341
PR	02-OCT-2000	2000US-02370342
PR	13-OCT-2000	2000US-02399397
PR	20-OCT-2000	2000US-02409600
PR	20-OCT-2000	2000US-02411785
PR	20-OCT-2000	2000US-02411786
PR	20-OCT-2000	2000US-02411787
PR	20-OCT-2000	2000US-02418088
PR	20-OCT-2000	2000US-02418089
PR	20-OCT-2000	2000US-02418266
PR	20-OCT-2000	2000US-02422211
PR	01-NOV-2000	2000US-02446177
PR	08-NOV-2000	2000US-02466474
PR	08-NOV-2000	2000US-02466475
PR	08-NOV-2000	2000US-02466476
PR	08-NOV-2000	2000US-02466477
PR	08-NOV-2000	2000US-02466478
PR	08-NOV-2000	2000US-02466533
PR	08-NOV-2000	2000US-02466524
PR	08-NOV-2000	2000US-02466525
PR	08-NOV-2000	2000US-02466526
PR	08-NOV-2000	2000US-02466527
PR	08-NOV-2000	2000US-02466528
PR	08-NOV-2000	2000US-02466529

PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249222.
 PR 17-NOV-2000; 2000US-0249223.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249225.
 PR 17-NOV-2000; 2000US-0249226.
 PR 17-NOV-2000; 2000US-0249227.
 PR 17-NOV-2000; 2000US-0249228.
 PR 17-NOV-2000; 2000US-0249229.
 PR 17-NOV-2000; 2000US-0249230.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Disclosure; SEQ ID NO 13188; 1701bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (ABH14678-ABH18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
 CC
 XX Sequence 11234 BP; 3094 A; 2417 C; 2869 G; 2854 T; 0 other;
 SO

Matches 1438; Conservative 0; Mismatches 950; Indels 301; Gaps 21;
 QY 7192 ACCGAGCGGGGCTGTGTGTCAGCGCTGTATCTTAGACCTTGGAGGCTGAGCAG 7251
 Db 4957 AACGAGCCAGGTCGGGTGCTGTATCTTAGACCTTGGAGGCTGAGCAG 4898
 QY 7252 GTGATTTGCTTGAAGCGGCTTGAAGCTTGAAGCCAGGCTTGGAGGCTGAGCAG 7311
 Db 4897 GTGATTCAC-----GAGTCAGAGAGATCGAGACCATGTGTGCTTGAAGCAGTGA 4849
 QY 7312 ACCTCATCTTCAAAAAATA-TAAAAATTAGTCGGGTGTGTATGTCCTTATATGTC 7370
 Db 4848 ACCCGCTCTTCAAAAAATAAAAAAATTAAGCCGGTGTGTGAGGAGTGTGTATGTC 4789
 QY 7371 CCATCTACTTCAAGGCTGAG-----CCCGAAGGTGAGGCTTCAAG 7412
 Db 4788 CAGCTACTTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4729
 QY 7413 TGAGCGGTATCGTGTATGTCAGCTCAGGCTGGGTGAGAGAGTGAACATGTCGAA 7472
 Db 4728 TGAGCGGAGTTGGCGCATGTCAGCTCAGGCGGGTGAAGAGCGAGATCCGCTCAAA 4669
 QY 7473 AAAAAAACAACAAAAACA-----AACAAAAACA 7503
 Db 4668 AAAAAAACAACAAAAACAACAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4609
 QY 7504 ACAAACAACAAAAACCAATATATATATATATATATATATATATATATATATATATAT 7563
 Db 4608 GACAGCAAGGAAATGTCCTTACTGCAATCTTGGCTGTGAATCAGGGGTGAGATCGAG 4549
 QY 7564 TTTTGTAAATATAC-----AACATGACC 7588
 Db 4548 TTGTGAAGACCTCCCGAGTATTCACAGGCTATGTCATGATTAATAGTTAAGTTT 4489
 QY 7589 CAGCTACAGATGAGGAGATCCCTCCCTCTCATCTGTAAATTTTCTTCTGACTCA 7648
 Db 4488 ATACTACCTGAATAGACAGAAACATTCAGCCCAAAATTTATTTATTTATTTATTC 4429
 QY 7649 CAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7708
 Db 4428 TTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 4369
 QY 7709 GAGTGCATGCGCAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7768
 Db 4368 GAGTGCATGCGCAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4309
 QY 7769 CTTGCTCAGCTTCCGATATGCTGGAGTACAGGCGCATACCATGCTGGCTTAAT 7828
 Db 4308 CTTGCTCAGCTTCCGATATGCTGGAGTACAGGCGCATACCATGCTGGCTTAAT 4249
 QY 7829 TTTGTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7888
 Db 4248 TTGTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4229
 QY 7889 AGAGATAGGTTTACATGTTTGGCCAGGCTGTCTGAACTCTGACCTGAGTATCC 7948
 Db 4228 -----GGTTTGGCAATGTTGGCCAGGCTGTCTGAACTCTGACCTGAGTATCC 4176
 QY 7949 GCTTGTCTGCTCCCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8008
 Db 4175 GCTTGTCTGCTCCCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8064
 QY 8009 CATAGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 8064
 Db 4115 CCAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4056
 QY 8065 ATCATGTTCTTCACTTCTTCACTTCTTCACTTCTTCACTTCTTCACTTCTTCACTTCTT 8124
 Db 4055 AGAAGAAATATGATATCTCTCTGAGAAAGTACCTTCAAGCAAACTTAAAAAATC 3996
 QY 8125 CCT--CCTCAGTGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8171
 Db 3995 CTTGAGTTCAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3936

Query Match 3.6%; Score 357; DB 22; Length 11234;
 Best Local Similarity 53.5%; Pred. No. 1.2e-58;

QY 8172 TCTGAATGCCCCAGTCCATATTTTTCAGACCAACCATTTAAAGTGTGATTGCCAGCC 8231
 DB 3935 TATGTGTCTCTAATAAATATCCACAAACAAAGACAGAGTAGGCCACCATTA 3876
 QY 8232 TATGTAATGGCTCAATTAATGGGTGGGAATCATCTATTTACTTTCATGCAAGAGC 8291
 DB 3875 TGAGAAATGACAAATATCATTAATAGAGAAATATGCCACAGAGCTTTGATATTGA 3816
 QY 8292 ATAGCTCTGTTTCAAAATAGGCCCCCTGGCCAGGTGTGTGCTCATGCTTATATC 8351
 DB 3815 ATATATTTGTGATATTAAGAAATATGTTTGCAGAGAGCGGTGGCTCAACCTGTATC 3756
 QY 8352 CCAACACTGGAGAGGCC-GAGGGGGGAGATCATCTTGAATCCAGAGTCTAGACAGC 8410
 DB 3755 CACGACTGTGGAGAACCGAGGTGTGTGATCATCTGAGGTCAAGAGTTGAGACAGC 3696
 QY 8411 CTGGGCAACATGATGAAATCTCATCTCTACTAATAAATACAAAAATTTAGCCAGGTGTG 8470
 DB 3695 CTGGCCAAATATGTTGAATCCCGTGTCCACAAAATAC-AAAAATTAGCCGGCGTGT 3637
 QY 8471 GCGATGACCAATGATCCAGCTGTTCAGAGGCTGAGGCTGAGAGATTCCTGAGTGTG 8530
 DB 3636 AGCAGCTGCTGTATGCTGTGGCTACCTGGAGGCTGAGGAGATCATCTTGAACCCG 3577
 QY 8531 GGAGGAGAGATTGCACTGAAACCGTGAATGCTGCTGCAATCCAGCTGGGTGACAGAT 8590
 DB 3576 GGAGGGAGAGTTGCACTGAAACAGATCATACCACTGCACTCAGCTGGGTGAGAGAG 3517
 QY 8591 TGAGACCTGTCTCAAAAAACAATTAATTAATTAATTAATGTTCTGAGCAGG 8650
 DB 3516 CAAGATCTCATCTCAAAAAACAATTAATTAATTAATTAATGTTCTGAGCAGG 8650
 QY 8651 TAAATTCAGTGGGAACTCTCCAGGGAGGTGATGTCAGTCCGCTGTAATCTCAG 8710
 DB 3474 -----ATATATTAAGCCCTGATTCAGACCG 3446
 QY 8711 TACAGGCTAATTAAGAACTTGTGTAGCAGAGAAACCTGATTTACTCAACAA 8770
 DB 3445 GACAGAAAATATATGATATCATAGGAATTAATCACTAATAATGACAGACTTCCATA 3386
 QY 8771 TATTTGTGACATCTGATTAAGAGTGGGCAATGCTCTAGGCACTGATACAGTATCA 8830
 DB 3385 CATTGCTGTGTTTAAAAAAA-----AAAAAAA 3348
 QY 8831 ACATGCAACAAAGATGCTGCTGCAAGCTCTGCTAAGTGAAGAGAGCAATTAAGA 8890
 DB 3347 ACAGGCCAGGTGGGTGCTCAC-----ACCTGTAATCCGGCATTG 3303
 QY 8891 AAGAGAAAGAGAAAGAAATTAATTTAGTAAATTAAGAGTGTAAAGAAATTAAGA 8950
 DB 3302 GAGGGCAAGGCAAGATCAGAGGTGAGAGATTTGAGAGTGGCTGGCCACATGTGTG 3243
 QY 8951 CAGGATAGTGGATAGAGGTGAGAGAAATGAGGGCTGTCTTGAAGAAATGATTTTGA 9010
 DB 3242 AAACCCGCTCTACTAATAAATACAAAAATCACTGAGGCTGATGAGCACTGCTGTAAT 3183
 QY 9011 GCTGAGACTTCAATGATGAGAGAAATTAACAACAGATGTGTGAGAGAAAGCAATTTT 9070
 DB 3182 CCCAGCTACTCGGAGGCTGAGGAGAGAGCTGTTGAATCGGGAGCA----- 3132
 QY 9071 AGGGAGGTGAGAGACATTAATCTTCAAGAAATCAAGAAAGAACCTGTGAGGCTGAAC 9130
 DB 3131 ---GAGGTGACAGTACCGAGATCAAGCATTAATTAACAGCTGGGTGACAGACAAAG 3075
 QY 9131 ACAGAGAAAGAGAGGTGGTGAATTAAGAGGCAAGGAGCGAGTGGCCAGGTACTTA 9190
 DB 3074 AATCTGACAAAAAAGATTTTAAAGAA-----CAATTTCTAGAA 3028
 QY 9191 GACCTGTAAAGGTTTCAACATTAAGAGAGTCAATCAGAAAGCTTGAAGAGGCTGT 9250
 DB 3027 ATGTAGATACAGTAATTAACAATAAATACTAATAATATATATGCTAT 2686

QY 9251 GATATATCTAATCAATTTTATTAAGATCACTGACCTTTTGCAGAAATTAAGTTA 9310
 DB 2967 GTGTGATGTGTATGATATATATATATATATCTGTAAGAAATTAAGGTGCAATTA 2908
 QY 9311 TAAAGTACAGCATGTAAGACAGAAATCAAGTACCAATCCGTGCAATTTCAATTA 9370
 DB 2907 ACATATCAAGAGGTGTAAGGACTGAAT-----TTGCTCCCTTCTGCTACCAAAAT 2852
 QY 9371 GAGGTATGACCGCTTGAAGTATGATGACAGAGAGTGTGAGAAATCAATGAT 9430
 DB 2851 CCTGTGTATAGTCTTAACCCCACT-----ACCTCAGCTGGAT 2811
 QY 9431 ATATTTGAGAGTGAAGTGAAGATTAATTAATGATAGATAGAGCCGGGTGTGTG 9490
 DB 2810 TGCATTTGAGAAAGG-----TCTTTAAGAGTAACTTAAGTTGCTGGGCACTGTG 2756
 QY 9491 CTACGCTGTATCTTACACATTTTGGAGGCCAAGGCAAGTGAATCACTGAGTCAAG 9550
 DB 2755 CTACGCTGTATCTTACACATTTTGGAGGCTGAGAGTGTGATCACTGAGTCAAG 2696
 QY 9551 AGTTGAGACCACTTGAACCAATGATGAACTGTCTTCTAATAATAC-AAAAAT 9609
 DB 2695 AGTTGAGACCACTTGAACCAATGATGAACTGTCTTCTAATAATAC-AAAAAT 2636
 QY 9610 AGCTGGAAATGGTGCACATGCTGTATCTGACCTTCTGAGGCTGAGCAGAGA 9669
 DB 2635 AGCAGGCAATGGGAGGATGCTGTATCTCAG-CTACTTGGAGGCTGAGCAGAGA 2577
 QY 9670 ATGCTTGAACCTGGAGGTGAATGTTGCAAGTGAAGCCGAGATTCACATTCCTCAG 9729
 DB 2576 ATTGCTTGAACCTGGAGGAGAGAGTTCAGATGAGCTGAGATTGCACTTCCTCAG 2517
 QY 9730 CCTGGGAAACAGAGTGAATCTGCTCTAATAATTAATGAATGAATGA 9778
 DB 2516 CCTGACTTAACAGAGTGAATCTCCTCAATCAAAAAAAA 2468

RESULT 7
 ABK22784
 ID ABK22784 standard; cDNA; 57273 BP.
 XX
 AC ABK22784;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human high bone mass (HBM) polynucleotide clone #7.
 XX
 KW Human; mouse; Zmax1; HBM; high bone mass gene; lipid regulation; stroke;
 KW lipid-associated condition; arteriosclerosis; cardiovascular disease; ss;
 KW osteoporosis; atherosclerosis; diabetic atherosclerosis; plaque build-up;
 KW neurovascular condition; wound healing; gene therapy; PCR primer; probe;
 KW bone development disorder; antiarteriosclerotic; cardiovascular;
 KW osteopathic; cerebroprotective.
 XX
 OS Homo sapiens.
 PN WO200192891-A2.
 PD 06-DEC-2001.
 XX
 PP 25-MAY-2001; 2001WO-US16946.
 XX
 PR 26-MAY-2000; 2000US-0578900.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 PA (UYCR-) UNIV CREIGHTON SCHOOL MEDICINE.
 XX
 PI Carulli JP, Little RD, Recker KR, Johnson ML;
 XX
 DR WPI; 2002-097784/13.
 XX
 PT Identifying molecules involved in lipid regulation, useful for
 PT diagnosing, treating or preventing e.g., arteriosclerosis, comprises

XX Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
 KW gene therapy; bone density modulation; bone strength; trabecular number;
 KW bone size; bone tissue connectivity; bone disease; osteoporosis;
 KW osteomalacia; rickets; Paget's disease; neoplasm of the bone; gene; ds.
 OS Homo sapiens.
 XX MO200292764-A2.
 XX 21-NOV-2002.
 XX 13-MAY-2002; 2002MO-US14876.
 XX 11-MAY-2001; 2001US-290071P.
 XX 17-MAY-2001; 2001US-291311P.
 PR 01-FEB-2002; 2002US-353058P.
 PR 04-MAR-2002; 2002US-361293P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA (AMHP) MYETH.
 XX Babil P, Bex FI, Yaworsky PJ, Bodine PV;
 DR WPI; 2003-129278/12.
 XX New transgenic animals (e.g. mice), useful as models for studying bone
 PT density modulation, developing drugs for treating or preventing bone
 PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
 PT reduced bone density -
 XX Example 2; Page 362-384; 603pp; English.
 XX The invention relates to novel transgenic animals expressing the high
 CC bone mass (HBM) gene, expressing the corresponding wild type HBM gene,
 CC comprising an alteration of the gene encoding LRP5 or LRP6, or
 CC expressing an LRP5 that is modulated by an altered gene control.
 CC sequence introduced by homologous or non-homologous recombination. The
 CC transgenic animals are for the study of bone density modulation or bone
 CC mass modulation. The invention has osteopathic and cytostatic activity.
 CC The polynucleotides of the invention may have a use in gene therapy.
 CC The transgenic animals and nucleic acids are for the study of
 CC bone density modulation, where the bone mass is modulated relative to
 CC non-transgenic animals of the same species in more than one parameter
 CC selected from bone density, bone strength, trabecular number, bone
 CC size, or bone tissue connectivity. The transgenic animals, nucleic
 CC acids and methods are useful for identifying molecules involved in bone
 CC development, and for developing pharmaceutical compositions, which may
 CC be employed for treating or preventing bone diseases, e.g.
 CC osteoporosis, osteomalacia, rickets, Paget's disease, or neoplasms of
 CC the bone. The transgenic animals and nucleic acids are also useful in
 CC methods for diagnosing diseases involved in bone development, or
 CC characterised by reduced bone density or mass. The present sequence is
 CC used in the exemplification of the invention.
 XX Sequence 66933 BP; 14237 A; 17817 C; 18323 G; 16556 T; 0 other;
 SQ
 Query Match 3.5%; Score 355; DB 25; Length 66933;
 Best Local Similarity 59.6%; Pred. No. 4.1e-58;
 Matches 871; Conservative 0; Mismatches 460; Indels 130; Gaps 11;

QY 7364 TATAG-----TCCCATCTACTGAGGCTGAGCCCGGAAGGTGAGGCTTCACTGAGCC 7418
 Db 61431 ACTTGGAGTCTGAGAGCATGAGTAATGCTTGAACATGAGGAGGAGGTGACGTAGGCC 61490
 QY 7419 GTGATGTGTACTGCACTCCAGCTGGGTGACAGAGTGAACATGCTCCAAAAAAC 7478
 Db 61491 GAAATGCGCATTTGACCTCCAGCTGGGCGACAGAGCGAGTCTGTCTCAAAAAAAA 61550
 QY 7479 AAAAAAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 7538
 Db 61551 AAAAAAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 61610
 QY 7539 ACCTAGCTGAGTGAAGATGCACTATTGTTGTAATATCAACCAATGACCCAGCTACAGC 7598
 Db 61611 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 61663
 QY 7599 ATGGGAGTCTCCCTCCCTCTACCTGGTAATTTTCTTCTGACTCACTATTTGTT 7658
 Db 61664 ---AGGCTTCCCTGCTGATCATAAAAACAACCTTAATTTTCACTCTCTCTTTT 61720
 QY 7659 GTTGTGTGCTGTGTTGAGATGAGTCACT-CTGTCAACCCAGCTGAGTGCAT 7717
 Db 61721 TT-----TTGAGCAAGATTGCTCTGTGCTCCAGCTGAGTGCAT 61766
 QY 7718 GAGGCAATCTTGGTCACTGCAACTCTGCTCTGAGTTCAAGCATCTCTGCTCTCA 7777
 Db 61767 GGGGCAATCTCACTCACTGTAACCTCCGCTCCGGGTGAGTGAATCTCTGCTCTCA 61826
 QY 7778 GCTTCCCTTACTGAGTCAAGGCGCATACCACTGATCCCTGCTTAATTTTGTATTT 7837
 Db 61827 CTTCCCGATGACTGGGATTAAGCATGACACCACTGCTGCTAA----- 61875
 QY 7838 TTTTGGGTTACAATGACTATTATTAATTTTGTATTTTGTATAGAGA-TAG 7896
 Db 61876 -----TTTGTACTTTAGTAGACGG 61899
 QY 7897 GATTTCACATGTTGCGCAGGCTGTCTGCACTCTGACCTGACGTATCGCTGCT 7956
 Db 61900 GATTTCACATGTTGCGCAGGCTGTCTGCACTCTGACCTGACGTATCGCTGCT 61959
 QY 7957 CGGCTTCCCAATGCTGATTAACAGGATGAGCAACCACTGCTGCTGCTGCTGCT 8016
 Db 61960 CGGCTTCCCAATGCTGATTAACAGGATGAGCAACCACTGCTGCTGCTGCTGCT 62019
 QY 8017 TTTATGATCTCTTGTGCTCTTCAACACTTGGCTGACGTCATCATGTTCTCT 8076
 Db 62020 TCTTAAGCTCTCTCTCTTCTGTTAAGTTCATCT-----T 62058
 QY 8077 CCACTTCTCACTTCACTGATGATCTTCACTGCTGCTGCTGCTGCTGCTGCT 8136
 Db 62059 TAAATCTTGGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 62118
 QY 8137 GCTCTTTTCTTCTGATGAATTTCCAGAGGGAATCTGAATGCGGCACTATTTTC 8196
 Db 62119 GCTGATGCTTAAAGTGGGCTGAATTAAGTGGGCTTATCAGAGGCAAGAACTCTATG 62178
 QY 8197 AGACACACCACTTAAGTGGTGAATGCGCAGCTATGATTTGGCTACTATATGGGTT 8256
 Db 62179 AG-----AATTTGAACAGAAAGTTCCTGCTACAGCTTATTAACAGGACGTGAA 62230
 QY 8257 GGGAACTCATCTTATCTTCACTTGAACAAAGCACTAGCTGTGTTCTCAAAATAGGGC 8316
 Db 62231 TACAGAAATTAAGACGTGATGTAACAGAACTCT-----AAGAAATGCA 62277
 QY 8317 CCTTGGGCGCAGGTGTGTGCTCATGCTTATATCCCAACATCTGTGGAGGCGGAGG 8376
 Db 62278 GAATAGCGCAGGATGTGTCTCAACCTGTCTCAACAGCACTTTGGGAACCAAGCGG 62337
 QY 8377 GCAATCACTTGAAGTCAAGAGTCTTCAAGCAGCTGCGGCAATGATGTAATCTCATCT 8436
 Db 62338 GTGATCTACTTGGGCTCAGAGGATTCAGACCACTGCGCAATATGTGAAACCCATCT 62397
 QY 8437 CTACTAAAAATAC-AAAAAATTAGCAAGTGTGTGATGACACAGTACTCCAGCTGT 8495

Db 62398 CTAATAAATACAAAAAATTAGCTGGTGGTGGATGCGATGCTTAATCCAGCTAC 62457
QY 8496 TCGAGAGGCTGAGGTGGAGAGATTGCTCGAGTGTGGAGCGAAGATTGAGTGAACCGT 8555
Db 62458 TCGGAGGCTGAGAGGAGAAATCACTTGAACCTGGAGCGAGGTTGACGTAGCCGA 62517
QY 8556 GACTGTGCTCTGCAATCCAGCCTGGGTGACAGATTGAGACCTGTCTCAAAAAACAAT 8615
Db 62518 GATCATGCCACTTACTCCAGCTGGTGGTGGAGAGCCGAACTGTGAAAAA 62577
QY 8616 AATTAATTAATTAATATATG 8636
Db 62578 AAAAAAGAGATTCAACTTG 62598

RESULT 10
AAL05708
ID AAL05708 standard; DNA; 27154 BP.
XX AAL05708;
AC
XX 21-NOV-2001 (first entry)
DT
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 8396.
XX
KM Human; reproductive system related antigen; reproductive system disorder;
KM cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234423.
PR 21-SEP-2000; 2000US-0234474.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.

QY 8542 TTGACGGAACCGGACCTGCTGCTGATCCAGCTGGGTGACAGATTGAGACCCCTGT 8601
 DB 1923 TTGACGTGACCTGATGATGAGCCACTGCTGCTGAGGCAACAGGAGACTTCAT 1982
 QY 8602 CTCAAAAACAAATTAATAATAATAATAATAT 8635
 DB 1983 CTCAAAAAAATTAATAATAATAATAATTAAT 2016

RESULT 11
 ABL45808
 ID ABL45808 standard; DNA; 9163 BP.
 XX ABL45808;
 AC
 XX 26-APR-2002 (first entry)
 DT
 XX
 DE Human endothelial differentiation, G-protein coupled receptor 6 gene.
 XX
 KM Human; endothelial differentiation, G-protein coupled receptor 6;
 KM EDG6; haplotype; cancer; angiogenesis; inflammation; chromosome 19p13.3;
 KM cytostatic; antiinflammatory; gene therapy; SNP;
 KM single nucleotide polymorphism; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT allele /tag= a
 FT allele /tag= a
 FT allele /tag= b
 FT allele /tag= c
 FT allele /tag= d
 FT allele /tag= e
 FT allele /tag= f
 FT allele /tag= g
 FT allele /tag= h
 FT allele /tag= i
 FT allele /tag= j
 FT allele /tag= k
 FT allele /tag= l
 FT allele /tag= m
 FT allele /tag= n
 FT allele /tag= o
 FT allele /tag= p
 FT allele /tag= q
 FT allele /tag= r
 FT allele /tag= s
 FT allele /tag= t
 FT allele /tag= u
 FT allele /tag= v

FT allele replace(5587,T)
 FT /tag= w
 FT allele replace(5606,C)
 FT /tag= x

PN MO200206446-A2.
 XX
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001WO-US22523.
 XX
 PR 17-JUL-2000; 2000US-218727P.
 XX
 PA (GENA-) GENAISANCE PHARM INC.
 XX
 PI Kitem SE, Koshy B;
 XX
 DR WPI; 2002-171804/22.
 DR P-PSDB; AAM48981.
 XX

PT New genetic variants of endothelial differentiation, G-protein coupled
 PT receptor-6 gene for studying expression, function of the gene and
 PT expressing EDG6 protein for use in screening drugs to treat cancer,
 PT inflammation -
 XX
 PS Claim 20; Fig 1; 11pp; English.
 XX
 CC The present invention provides the gene, protein and cDNA sequences of
 CC the human endothelial differentiation, G-protein coupled receptor 6
 CC (EDG6). Also identified are single nucleotide polymorphisms (SNPs) found
 CC within the sequences. The sequences can be used in the identification of
 CC the haplotype of an individual, and in the treatment of cancer,
 CC angiogenesis and inflammation. The present sequence is the EDG6 gene,
 CC which is found on chromosome 19p13.3.
 XX
 SQ Sequence 9163 BP; 1730 A; 2757 C; 2622 G; 2054 T; 0 other;

Query Match 3.5%; Score 346.2; DB 24; Length 9163;
 Best Local Similarity 57.3%; Pred. No. 1.4e-56;
 Matches 847; Conservative 0; Mismatches 573; Indels 57; Gaps 10;

QY 7179 TATATTAATAAATATACCGAGCGGCGGTGAGTACCGCTGTATCTAGACCTTTGG 7238
 DB 7442 TATTTTAAGAGATGAGGCGCGGCGGTGAGTACCGCTGTATCTAGACCTTTGG 7501
 QY 7239 GAGGCTGAGGAGGTGATGCTTGAAGCGGCTTGAAGATTGAAGACCGCTGG 7298
 DB 7502 GAGGCGGAGGTGAGGCGGATCAC-----GAGGTGAGGAGATGAGACCATCTGG 7550
 QY 7299 GCACATGCGAAACCTCATCTCTACAAAATATATAA-ATTATGCGGTGTGTAGTG 7357
 DB 7551 CTAAACAGGTGAAGCCCATCTTTACTAAATAATACAAAGATTAGCCAGCGGTGTGGG 7610
 QY 7358 CGTTCTATAGTCCCATCTACTTACAGGCTGAG-----CCCGAAG 7399
 DB 7611 GGTGCTGTAGTCCAGCTACTCTGAGGCTGAGTACGAGATGCGGTGAACCCGGAG 7670
 QY 7400 GTGAGGCTTCAAGTACCGGTGATGCTACTGCACTCAAGCTGGGTGACAGATGAG 7459
 DB 7671 GCAGAACTTGAGTACGCGAGATGAGGCACTGCACTCAAGCTGGGTGACAGATGAG 7730
 QY 7460 ACCATGCTCAAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 7519
 DB 7731 ACTCCGTCTCAAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 7790
 QY 7520 CCCATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7579
 DB 7791 GCCCAGGCTGTCTCAAACTCTG-----GCTCAAGTATCTCTCCACCTGACCTTCT 7845
 QY 7580 AACATGACCAAGTACAGATGAGGAGTCCCTCCCTCTCACTGTAATTTTCTTTTC 7639
 DB 7846 GAATAGCTGAGACTACAGAGATGACCAACCAAGCTGCTCATTTTCTATTTTTCATTT 7905

PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234224.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249287.
 PR 17-NOV-2000; 2000US-0249289.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251989.
 PR 06-DEC-2000; 2000US-0251479.
 PR 06-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM,
 XX
 DR WPI; 2001-457726/49.
 XX
 PT Isolated polypeptide for treating, preventing and prognosing disorders
 PT related to the endocrine system including endocrine disorders,
 PT reproductive disorders, and gastrointestinal disorders and also for
 PT testing and detection e.g. diagnosis -
 XX
 PS Disclosure; SEQ ID No 470; 558pp; English.
 XX
 CC The invention relates to cDNAs encoding novel human endocrine
 CC antigens or a fragment having biological activity, a domain, an epitope,
 CC full length protein, variant, allelic variant or a species homologue of
 CC the cDNA/antigen. The DNAs and polypeptides are useful for preventing,
 CC treating or ameliorating a medical condition when administered
 CC (e.g. by gene therapy or antisense-therapy). Identifying mutations in
 CC the genes coding for the antigens is useful for diagnosing a pathological
 CC condition or a susceptibility to a pathological condition. The DNAs,
 CC antigens and antibodies raised against the antigens useful for treating,
 CC preventing and/or prognosing disorders related to the endocrine system
 CC or hormone imbalance or reproductive disorders, cancers of endocrine
 CC tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal
 CC glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the
 CC hypothalamus and testes (e.g. vanishing testes syndrome), many examples
 CC of diseases and disorders are given in the specification. The present
 CC sequence is genomic DNA fragment form a gene encoding an endocrine
 CC antigen of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 29543 BP; 6367 A; 8023 C; 7634 G; 7519 T; 0 other;
 Query Match 3.4%; Score 343.8; DB 22; Length 29543;
 Best Local Similarity 58.3%; Pred. No. 4.9e-56;
 Matches 878; Conservative 0; Mismatches 482; Indels 145; Gaps 10;
 QY 7198 GCCGAGCTGTGTAGTACGCGCTGTATCTTGTGAGAGTGTGAGCGAGTGTGAT 7257
 DB 5325 GCCAGGCACAGTGTGCTACGCGCTTATACCTTACCGCTTTGGAGGTCAAGGAGAGT 5384
 QY 7258 TGCCTGAGCGGCTTGTAGCTTGTAGACCAAGCTTGGGCAATGCAAAACCTCA 7317
 DB 5385 TGCCTGAG-----CCAAGAGTTTGAACCAAGCTTGGGCAATGCAAAACCTCA 7317
 QY 7318 TCTCTACAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7377
 DB 5436 TCTCTACAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5495
 QY 7378 CTTCAAGAGCTGAGCCCGGAAGTGC-----AGCTTCAAGTGAAGCCG 7419
 DB 5496 CTGGGAGGCTGAGCGCGGAGATGTGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 5555
 QY 7420 TGAATGTGTACTGACATCTCAAGCTTGGGTGACAGATGACCAATGTCTCAAAAAACA 7479
 DB 5556 TGAATGTGTACTGACATCTCAAGCTTGGGTGACAGATGACCAATGTCTCAAAAAACA 7515

OY	7480	AAAACAAAAACAAAACAAA-----ACAAAACAAAACAAAACAAA	7517
Db	5616	AATGCAAGCTGAAGAGCCCAAGTCCATCTGGAGCGTCGATCTGGAAATGCAATCACAG	5675
OY	7518	AACCATATATATATATATATATACCTAGCTGAGTGAAGATGACATATTTTGGTAAATCA	7577
Db	5676	ACTTCTGACAGCTCATCTCAGAGGCCACAGAGGCTGGATTAGGGTAAGAAACAAAGTAAGCA	5735
OY	7578	CCAACTAGACCAAGCTACAGCATGGGGCAGTCCCTCCCTCTCACTGGTAAATTTTCTT	7637
Db	5736	CTCGATTAGGGTATTAATAATCTAAGGGGACACAGACAAATCTAAGGGAGACACAGACAA	5795
OY	7638	TCTGCACTCAAGTTTGTGT-----TGTGTGTGTGTGTGTGA	7681
Db	5796	ATTCACTTATCAAGATTAATATATATGTTTAAATGCAATGTGTGTGTGTGTGTGTGA	5855
OY	7682	TGAGATCTCACTC-TGTCACCCAGGCTGAGATGCAATGGCCCAATCTTGTTCACTGCA	7740
Db	5856	TGAGATTCACTCTGTGTGCCCCAGGCTGAGATGCAATGGCCGATCTCAACTCACTGAA	5915
OY	7741	CCTGTGCTCCTGGGTTCAAGGATCTTCCTGCTCAGCTCCTCCGTATAGCTGGAGTAC	7800
Db	5916	CCCTCGTCTCCCGGGTTCAAGCGATCTCTGCTCAGCTCCTCCGAGTAGCTGGGATTAC	5975
OY	7801	AGGGCATACACACCATGCTGGCTAATTTTGTATTTTTTGGGTACATGTACAT	7866
Db	5976	AGGTGTGGCCACTAATGCCCAGCTAAT-----60030	
OY	7861	TTATTAATTAATTTTGTATTTTGTATAGATAGAGTTTCAACATGTCGACAGCTG	7920
Db	6004	-----TTTTTGTATTTTATAGTAAGATGGGGTTTCTCATGTGTTCACAGCTG	6051
OY	7921	GCTCGAACTCTGACCTCAGGTGATCCGCTGCTCGGCTCCCAAAATGTCTAGATTA	7980
Db	6052	GCTTGAACTCTGACCTAAGGTGATTCACCTGCTTGGCTCCCAAAAGTCTGAGATTA	6111
OY	7981	CAGCATAGCAACAGCGCTGGGCCCTCATAAGTTTT--ATCATCTCTTGCTCT	8038
Db	6112	CAGGATAGCAACACCGCGGCTCGCAATATGTTTTTAAAAAATCAAAATTAATGCA	6171
OY	8039	TCACAACTTGGGCTGACAGTGAACATCATGTTCTCTCACTTCTCACTACTTCATGA	8098
Db	6172	AAAAAATCCAGATGAACAAAATATCAAAATTTTAAATTAAGATCAGATGAATGAGG	6231
OY	8099	TCTTCACTCACTTCCAACTGATACCTCCCTCAGTGTCTTTTTTCTAGTAAGAT	8158
Db	6232	AAATGAAAAATCAGCAACCTCTTAAATGTCGTGATTTTAAATGTTGATTT	6291
OY	8159	TCCAGAGGGAAATGATGATGGCCAGTCCATAT-----TTAGACCAACACATTA	8212
Db	6292	TTTTTCCAGAGATTAAGTATGTTAAAAAATATTTTAACAATGAAGAAATGATGTATTA	6351
OY	8213	AAGTGTGATTTGCGAGGCTATGATTTGCTACATTAATGAGGTGGGAATCAT-----CA	8268
Db	6352	AAAGTCATTAATTTAAGTGAATGCTTTATTTATCAAAACAGAACTTATTAATCT	6411
OY	8269	TTTACTTCAATGCAAAAGCAGCATAGCTGTGTTCTCAAAATAGGCCCCCTGGCCAG	8328
Db	6412	TTTGCTTCAATGTAATGTAATTAAGAAATTAATGAACCGGGCAGAGTGGCTGGCCGG	6471
OY	8329	TGTGTGCTCATGCTCTAATCCCAACATGTGGGAGGCCGAGGGGGCAGATCATTTG	8388
Db	6472	TGTGTGCTCATGCTCTAATCCCAACATTTGGGAGGGTGAACCAAGTGAATCATTTA	6531
OY	8389	AGTCAGAGTTCTTGACCAAGCTGGGCAACATGGTGAATCTCATCTCTAATAAATA	8448
Db	6532	AGTTCAGAGTTCAAGACCAAGCTGGCCAACTGTGTAAACCTGTCTCTATAAAAAA	6591
OY	8449	CAAAAAATTAGCCAGGTGTGTGTGATGACCAAGTATGCCAGCTGTTCAAGAGGCTAG	8508
Db	6592	CAAAAAATTAGC-----AGCGAGGGGTGCTCGGGAAGCTAAG	6628
OY	8509	GTGGAGGATTTGCTGAGTGTGGAGGACAGATTTCACTGAACCTGATCTGTGCTGTG	8568

Accession	Sequence	Position
Db	6629 ACAGAGAAATACCTTGAACCGAGGAGGTGGACGTAGGCGAGATCGTCCACAG	6688
Qy	8569 CAATCCAGCCTGGGTGACAGATTGAGACCTGTCTCAAAAAACAATTAATTAATAA	8628
Db	6689 CCCCTCAGCCTGGGTGACAGAGTAGGCGTCATCTCAAAAATAAAAAATAAACATAAAAA	6748
Qy	8629 TAAAT	8633
Db	6749 TTATT	6753

RESULT 13

ABT10147/c
ID ABT10147 standard; cDNA; 185035 BP.

AC ABT10147;

DT 04-DEC-2002 (first entry)

DE Human breast cancer associated coding sequence SEQ ID NO: 281.

KW Human; breast specific gene; breast cancer; differential expression;

22

XX
XX
.....

2000

0
1
4
7
8
9
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PR 25-APR-2001; 2001US-286090P.

XX
(Contd.) Contd. Page 27

2005

XX 2000 574003/75

XX
XX

PT gene expression in cell or tissue samples, where a differential gene

[illegible]

XX
XX
3

CC patient, which comprise detecting the level of expression in a tissue sample of two or more genes selected from those shown in AB009867-

CC ABT11112, where a differential expression of the genes indicates breast

CC progression, and in monitoring treatment of breast cancer in patients.

the onset or progression of breast cancer. The breast cancer genes may be used as diagnostic markers for the prediction or identification of the

CC malignant state of breast tissue, for confirming the type and progression

CC coding sequence of the invention.

CC specification, but was obtained in electronic format directly from WLPD
CC at etn.int@wlpd.nl and published not reviewed

XX
XX
COMM: 19603E BD- 433EC A. E17337 C. E1310 C. 30843 T. 0 Other.

3 48. 66666 343 8. NB 34. Fourth 19E02E

Best Local Similarity 58.3%; Pred. No. 6.9e-56;

[illegible]

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

Db	37219	GCCAGGCAACAGTGGCTCACGCGCTATTAACCTTCAGGCGTTTGGAGGTCACAGCAGAGGANT	37166
Qy	7258	TGCTTGAAGCGGCTTGACGCTTAGAGAGTTTGAGACCAAGCTTGGCAACATGGCAAACTCTCA	7317
Db	37159	TGCTTTGAGAG-----CCAAAGATTGAGACCAAGCTTGGCCAACTAGCAAGACCCCTG	37109
Qy	7318	TCTCTACAAAATAATATATAAATTAGTGGGGTGTTGTCGTTCCATATGTCCTCATTA	7377
Db	37108	TCTCTACTAAATAATATAAGATTAGCCGGGGAATGGTGTGCATGCTGTATGCTCCAGCTA	37049
Qy	7378	CTTCAGAGGCTGAGCCCGGGAAGTGC-----AGCTTCAGTGAAGCCG	7419
Db	37048	CTGGGGAGGCTGAGCGCGGAGAGATGGCTTGAGAGCTTGAGAGTTTGAAGGCTCCGCTGAGCTG	36989
Qy	7420	TGATCGTGCTATCTGCACTCCAGCGCTGGGTACAAGATGAGACATGTTCTCAAAAAAACA	7479
Db	36988	TGATGGTGCCACTCATTTCCAGCATGGGTGACAGATGAGACACTGCTCTAAATTAATA	36929
Qy	7480	AAACCAAAAAACAACAA-----ACAAACAACAACAACA	7517
Db	36928	AATCAAGACTGAAGAGCCCAAGTGCATCTGGGCGTGCATCTGGAATGCAATCACAG	36869
Qy	7518	AACCATATATATATATATATATACCTAGCTGAGGTAGAAATGACATAATTTGGTAAATACA	7577
Db	36868	ACTTTCTGCAAGCTCATCTCAGGCGCACAGGCGCTGGATTAGGCTGAACAAAGTAAGCA	36808
Qy	7578	CCAACATGACCCAGCTACAGCATGGGGCAGTCCCTCCCTCTCACTGTAAATTTTCTT	7637
Db	36808	CTCGCTTAGGGTATTAATCTTAAGGGACACAGACAAATTTCAAGGGGCAACAGACAA	36749
Qy	7638	TCTCTGACTCAGATTTTGTGT-----TGTTGTGCTGTGTTTGAACA	7681
Db	36748	ATTAGTTATCAAAATAAATGAATGTTTAAATGCAATGTGTTTGTGTTTGTGTTTGGACA	36688
Qy	7682	TGAGTGTCACTC-TGTACCCAGGCTGAGGTGAATGGGCAATCTGGTTCATCTGCAA	7740
Db	36688	TGAGTGTCACTCTTGTGCGCCAGGCTGAGGTGAATGGGTATCTCAACTCACTGAAA	36629
Qy	7741	CCTCTGCCTCTGGGTTCAAGCAATCTCTGCTCAGGCTCCCGTATAGCTGGAGCTAC	7800
Db	36628	CCTCCGTCCTCCGGGTTCAAGCAATCTCTGCTCAGGCTCCCGATAGCTGGAGTTAC	36568
Qy	7801	AGGCGCATACCACATGCTGCGCTAATTTTGTATTTTTTTGGGTTACATATATAT	7860
Db	36568	AGGTGTGCGCACATATGCCAGCTAAT-----	36541
Qy	7861	TTATTAATTTAATTTTGTATTTTATAGATAGGGTTTCCACATGTGGCAGGCTG	7920
Db	36540	-----TTTTGTATTTTATATAGATAGGGGTTTCCATGTTGTCCAGGCTG	36493
Qy	7921	GTCTCGAATCTTGACCTCAGGTGATCCGCTGCTCGGCTCCCAAAGTGTAGGATTA	7980
Db	36492	GTCTTGAATCTTGACCTTAAGGTATCTACCTGCTTGCGCTCCAAAGTGTAGGATTA	36433
Qy	7881	CAGGCTAGAGCAACACGCGCTGCGCTCATAGGTTTT--ATCTATCTCTTGGCTTCT	8038
Db	36432	CAGGCTAGAGCAACACGCGCGCTGCTCATATGTTTTTAAAAAATTCAAAATTATATGCA	36373
Qy	8039	TCACAATCTTGGGCTGACGATGACCAATATGTTCTCTCCACTTCTCACTACTTCATGA	8098
Db	36372	AAAAAATCCACAAATGAACAAAATATCCAAATTTTAAATTAAGATCAGATGAACGTAGGG	36313
Qy	8099	TCTTTCACTCAGTTCCAACTGATACCTCCCTCAGTGTCTCTTTTCCATAGATTA	8158
Db	36312	AAATGAAAAACATGCAACCTCTTAACATGTTCTGTGTGCAATTTTTTAATGGTTGATTT	36253
Qy	8159	TCCAGAGAGGAATCTGAATGGCCCACTCATAT-----TTCAGACCAACCAATTA	8212
Db	36252	TTTTTCCAGAGGAATTAAGTATTAATAAAATATTTTAAACAATGAGATGTATGATTA	36193
Qy	8213	AAGTGTTGATGTGCAAGCTATGATATGGTATCAATTAATGGGTGGAACTCAT---CA	8268
Db	36192	AAAGTCAATCAATTTTAAAGTAAATGCTTATTTATTCAAAACGAACCTTATTTATCT	36133

[illegible]

PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226661.
 PR 22-AUG-2000; 2000US-0236868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 03-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 06-DEC-2000; 2000US-0251856.
 PR 06-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-457726/49.
 Isolated polypeptide for treating, preventing and prognosing disorders related to the endocrine system including endocrine disorders, reproductive disorders, and gastrointestinal disorders and also for testing and detection e.g. diagnosis -
 PS Disclosure; SEQ ID No 471; 558bp; English.
 XX
 PS The invention relates to cDNAs encoding novel human endocrine
 CC antigens or a fragment having biological activity, a domain, an epitope,
 CC full length protein, variant, allelic variant or a species homologue of
 CC the cDNA/antigen. The DNAs and polypeptides are useful for preventing,
 CC treating or ameliorating a medical condition when administered
 CC (e.g. by gene therapy or antisense-therapy). Identifying mutations in
 CC the genes coding for the antigens is useful for diagnosing a pathological
 CC condition or a susceptibility to a pathological condition. The DNAs,
 CC antigens and antibodies raised against the antigens useful for treating,
 CC preventing and/or prognosing disorders related to the endocrine system
 CC or hormone imbalance or reproductive disorders, cancers of endocrine
 CC tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal
 CC glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the
 CC hypothalamus and testes (e.g. vanishing testes syndrome), many examples

of diseases and disorders are given in the specification. The present CC sequence is genomic DNA fragment form a gene encoding an endocrine CC antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 29521 BP; 6369 A; 8019 C; 7628 G; 7505 T; 0 other;

Query Match 3.4%; Score 342.2; DB 22; Length 29521; Best Local Similarity 58.3%; Pred. No. 1e-55; Matches 877; Conservative 0; Mismatches 483; Indels 145; Gaps 10;

```

QY 7198 GCGGGCGGTGAGTCAAGCCCTGTAATCTTACACTTTGGGAGCTGAGCGGTGAT 7257
DB 5325 GCGAGGCAAGTGGCTCAGCCCTATTAACCTCAGCGCTTGGGAGGTCAAGCGAGGAT 5384
QY 7258 TGGTTAGCGGCTTGAGCCTAGAGTTTGAGACCAAGCTGGGCAATGGCAAACTCA 7317
DB 5385 TGGTTAG-----CCAGAGTTTGAGACCAAGCTGGGCAATGGCAAACTCA 5435
QY 7318 TCTCTACAAATAATATAAATTAGTGGGTGTGTAGTGCCTTATAGTCCATCTA 7377
DB 5436 TCTCTACAAATAAATAAATAAATTAGCGGGATGTGTGTGCATGTTGTATGTCAGCTA 5495
QY 7378 CTTCAGAGGTGAGCCCGGAGGTG-----AGGCTTCAGTGAGCCG 7419
DB 5496 CTGGGAGGCTGAGCGCGGAGATGTGTAGCCTGGGATTTGAAGGCTCGCGGTGAGCTG 5555
QY 7420 TGAATGCTACTGACCTCAGCGCTGGGTGACAGAGTGAACCATGCTCAAAAAACA 7479
DB 5556 TGAATGCTACTGACCTCAGCGCTGGGTGACAGAGTGAACCATGCTCTTAATAATA 5615
QY 7480 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7517
DB 5616 AATGCAAGGTGAAGAGCCGAGTGCATCTGGCGCTGCATCTGMAATGACATCAACAG 5675
QY 7518 AACCCATATATATATATATATATATATATATATATATATATATATATATATAT 7577
DB 5676 ACTTTCTGAGCTCTCATCTCAGCCACACAGGCTGGGATTAAGGATGAACATGAAGCA 5735
QY 7578 CCAACATGACCCAGCTACAGCATGGGGAGTCCCTCCCTCAGCTGATTAATTTTCTT 7637
DB 5736 CTGGCTTAGGGTATTAATAATCTAAGGGGACACAGACAAATCTAAGGGGACACAGCAA 5795
QY 7638 TCTCTGACTCACAGTTTGTGT-----TGTGTGCTGTGTGTGAGA 7681
DB 5796 ATTCACTTATCAAGATTAATGATGTTTATGCAATGTGTGTGTGTGTGTGTGTGAGA 5855
QY 7682 TGAAGTCTCAGTC-TGTCAACCCAGGCTGAGTGCATGGCGCAATCTTGTGCTCACTGCAA 7740
DB 5856 TGAAGTCTCAGTCCTGTGCGCCAGGCTGAGTGCATGGCGTATCTCAATCACTGAGAA 5915
QY 7741 CCTCTGCTCTCTGGTTCAAGCATCTCTGCTCAGCCTCCCGATATAGCTGGGACTAC 7800
DB 5916 CTTCTGCTCTCTGGTTCAAGCATCTCTGCTCAGCCTCCCGATATAGCTGGGACTAC 5975
QY 7801 AGGCGCATACCAACATGCGCTGCTTAATTTTGTATTTTGTGGTTACATATATAT 7860
DB 5976 AGGTGTGCGCATATAGCCCAAGTAAAT-----6003
QY 7861 TTAATTAATTAATTTTGTATTTTGTATAGATATAGGCTTCAACATGTGGCCAGGCTG 7920
DB 6004 -----TTTGTATTTTGTATAGATATAGGCTTCAACATGTGTGCGAGCTG 6051
QY 7921 GTCTCAAGCTCAAGCTCAAGGATCCGCTGCTGCGCTCCCAAGGCTAGATTA 7980
DB 6052 GTCTCAAGCTCAAGCTCAAGGATCCGCTGCTGCGCTCCCAAGGCTAGATTA 6111
QY 7981 CAGGCAATGACCAACAGCTGCGCCCTCATAGGTTTT--ATCTATCTCTTGTGCTCT 8038
DB 6112 CAGGCAATGACCAACAGCTGCGCCCTCATATGTTTTTAAAAAATCAAAATTAATGCA 6171

```

```

QY 8039 TCACAACCTTGGCTTGACAGTGGAGCCATCATGTTCTCCCACTTCTCACTTATGA 8098
DB 6172 AAAAAATCCACAATGACAAATAATATCAAAATTTAATAAAGTACGATGAAGTGAAGG 6231
QY 8099 TCTTTCACTGCTAGTTCCAACTGATATCTCCCTCAGGTTCTCTTTTCTTCTAGTAAGT 8158
DB 6232 AATGGAAGAAATGACAGCCCTTAACATGTTGCTGTGATCTTTTATATGTTGATTT 6291
QY 8159 TCCAGAGAGGAATCTGAATGGCCCAATTTTTCAGACCA-----CACCAATT 8211
DB 6292 TTTTTCAGAGGATTAAGTAAATTAATAAATAATTTTACAAATGACAGATGTATTA 6351
QY 8212 AAAGTGTGATTTGCA---GCTATGTAATGGCTCATTAATAGGTTGGGAATCATCA 8268
DB 6352 AATGATATACATTTCTAAGTAAATGCTTTTATTTACCAAAACAAATTAATTAAT 6411
QY 8269 TTTACTTCAATGACAAACAGCATAGCTCTGTTCTCAAAATAGGCCCCCTGGCCAGG 8328
DB 6412 TTTGCTTCAATGATGAATGAATTAAGAAATTAAGTTAAGACCGGGGAGTGGCTGGCCGG 6471
QY 8329 TGTGTGCTCATGCTTATTAATCCCAACTGTGGAGGCCAGGGGCGCATCACTTG 8388
DB 6472 TGTGTGCTCATGCTTATTAATCCCAACTTTGGAGGGTGAAGCATGATGATCACTTA 6531
QY 8389 AGTCCAGAGTCTTGAACAGGCTGGGCAACATGGTAATCTCATCTTCAATAAATA 8448
DB 6532 AGGTCAAGAGTTCAACACAGGCTGGCCCAACATGGTAATCTCTCTCAATAAATAA 6591
QY 8449 CAAAAAATTAAGCAGAGTGTGTGTCATGCAAGTATGCCAGCTGTTGAGAGGCTGAG 8508
DB 6592 CAAAAATTAGCC-----AGGGAGGCTCTCGGAGGCTTAAG 6628
QY 8509 GTGGAGGATTTGCTGAGTGTGGAGGCAAGATTCAGATGAACCTGACTGTGCTCTG 8568
DB 6629 ACAGAGAAATCACTTGAACAGGAGGTGAGGTTGCAATGAGCCAGATCGGCCACAG 6688
QY 8569 CAATCCAGCTGGGTGACAGATTTGAGACCTGTCTCAAAAAAATAAATAAATAAATAA 8628
DB 6689 CCTTCAGCTGGGTGACAGAGTGAAGCTTCATCTCAAAAAATAAATAAATAAATAA 6748
QY 8629 TAAAT 8633
DB 6749 TTAAT 6753

```

RESULT 15
ABA20494/c
ID ABA20494 standard; DNA; 16877 BP.

ABA20494;
23-JAN-2002 (first entry)

Human nervous system related polynucleotide seq ID NO 12825.

Human; nootropic; neuroprotective; cyrostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary; antiParkinsonian; antistroke; antianemic; antiarthritic; cancer; antidiabetic; hepatotropic; cerebroprotective; antiinflammatory; antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal; antiparasitic; cardiac; immune disorder; cardiovascular disorder; neurological disease; infection; nephroretropic; gene therapy; vaccine; da.

Homo sapiens.
WO200159063-A2.
16-AUG-2001.
17-JAN-2001; 2001WO-US01334.
31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216860.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226681.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250390.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0251997.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX

PS Disclosure; SEQ ID NO 12825; 1701bp + Sequence Listing; English.

XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or uterine; (b)
CC immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

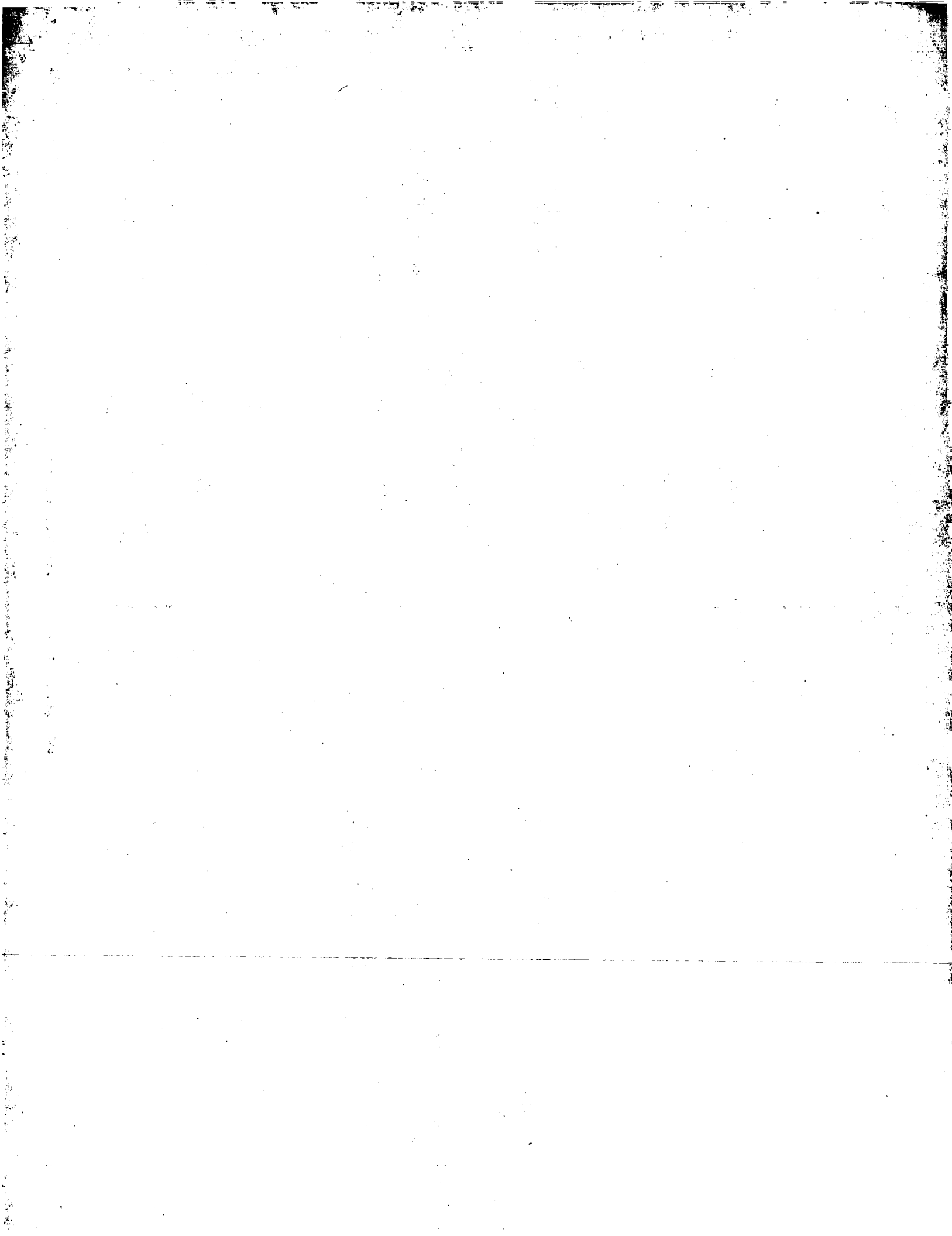
XX Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other:

Query Match 3.4%; Score 341.4; DB 22; Length 16877;
Best Local Similarity 58.5%; Pred. No. 1.3e-55;
Matches 876; Conservative 0; Mismatches 516; Indels 105; Gaps 12;

QY 7177 CCTATATATTAATAATCCAGAGCGGCGGTGTGAGTACAGCTGTATATCTAGCACTTT 7236
DB 10860 CGTAATGTGAAAGACATTTGGCCGGGCGGTGTGCTCACACCTGTATATCTAGCACTTT 10801
QY 7237 GGGAGGCTAGAGGAGGTGATGCTTGAAGCGGCTAGAGCTTGAAGTTGAGACAGCCCT 7296
DB 10800 GGGATTTCTGAGGAGGCGCAT-----CATGAGGTCAAGAGTCAAGACATCTCT 10752
QY 7297 GGGCAATGCGCAAAACCTCATCTCTACAAAATAA-TAAATAATGAGCGGTGTGAG 7355
DB 10751 GGTATATAGGTGAAAGCCCATCTCTATTAATAATCAAAAATTTAGCCGGGTGTGAG 10692
QY 7356 TGGCTTCTATAGTCCCATCTTCTTACAGAGCTGAGCCGGAAG-----GTGAG 7405
DB 10691 CGGGTCTGTATAGTCCCATCTTCTGAGGCTGAGGCAATATGCGTGAACCCGGAGG 10632
QY 7406 GCTTCAAGAGCGGTATGCTGCTACAGCTCAGCGTGGGAGACAGATGAGCAATG 7465
DB 10631 TGAATGTAAGCCAGATCAAGCACTGCACTCAGCTGAGTGAAGACGAGATCCCA 10572
QY 7466 TCTCAAAAAAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 7525
DB 10571 TCTCAAAAAA--AAAAAAGAAAAAGAAAGCACTTAATGCACTTAATCTAGAACATC 10514
QY 7526 TATATATATATATATACCTAGCTGAGGTGAGATGCACTATTTGGTAAATCACCAATG 7585
DB 10513 ATCACTTAGCTAGCTTAACTTAAATGTGCTCAGAAACATTTATTAAGCTTAC-----G 10460
QY 7586 ACCCACTACAGATGAGGAGATCCCTCCCTCTCACTGTAAATTTTCTCTCTGAC 7645
DB 10459 GCTGGGCAAAATATATAGCAAAAGCTATTTTCACTTATTAATGAAGAAAGTTGAAT 10400
QY 7646 TCACAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7705
DB 10399 CCGATGTATTTT-----TTTTTGAAGACAGTCTCACTGTGTGTGTGTGTGTGTGTGT 10350
QY 7706 CTGAGGTGCAATGAGCGCAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7765
DB 10349 CTGAGGTGCAATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10291
QY 7766 CCGTCTGCTCAAGCTCCCTGTATAGCTGAGATGCAAGGCGCATACCACTGCTGTGCTA 7825
DB 10290 TCTCCGCTCAAGCTCCCTGTATAGCTGAGATGCAAGGCGCTGCTCATATGCAACGAT 10231
QY 7826 ATTTGTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7885

DB 10230 AATTTTTTTTT-----TTTTGTATTTT 10206
QY 7886 AGTAGAGATAGGTTTACCATGTTGGCCAGGCTGTGCACTCCGACCTGACAGTGA 7945
DB 10205 AGTAGAGATAGGTTTACCATGTTGGCCAGGCTGTGCACTCCGACCTGACAGTGA 10146
QY 7946 TCCGCTGCTGAGCTCCCAAGTGTGAGTTACAGGATGAGCAACAGCCCTGAGCC 8005
DB 10145 TCTTCCGCTGAGCTCCCAAGTGTGAGTTACAGGATGAGCAACAGCCCTGAGCC 10086
QY 8006 C-----CTCATAGTTTATTTATCTATTTCTTGTCTTTCACAACTTTGCTTGCAGTG 8060
DB 10085 GAATCTCATGATATTTATTTAGATGCTTGTGCAAGTGAAGAAACAGAAATGTTGAGGG 10026
QY 8061 GACATCATGTTCTCTCCACTTTCACATCTTCACTTCACTTCACTTCACTTCACTT 8120
DB 10025 TATGGAGACAGTTCCCACTGAATATGATCATCTTCTCACCATCATAAAGTTCTTTT 9966
QY 8121 GATACCTCCCTGAGTTGCTTTTCTTCTATGATTTCCAGAGAGGAAATCTGAATG 8180
DB 9965 GGGCGGAGGAGAGGAGGAGGAGAGAGTATAGAAATGACCTGCACTGTGCTGTGGG 9906
QY 8181 CCCAGTCCATATTTTCAACCAACCACTTAAAGTGTGATTTGCGACCTATG----- 8235
DB 9905 CTCAGTTTCAACACTTAAAGCAACAGTCACTTCTGCAAGAACTGAACTCCGTGACCC 9846
QY 8236 --TATTGGTACATTAATGAGTTGGAACTCATCTTATCTTATGCAACAAAGCAAT 8293
DB 9845 CAATATCCCTTACTTACTGTAATTAAGCCCTGCTCTATGTTATAGGCTACTGA 9786
QY 8294 AGCTGCTTCTC-----AAATATAGGAGCCCTGAGGCAAGTGTGCT 8338
DB 9785 AGGATGTATAGCTTATATTAAGATTAAGAAATGAAACAGGCTGGGCGAGTGGCT 9726
QY 8339 CATGCTATTAATCCCAACCTGTGGAGGCGAGGAGGCAATCACTTGAATCCAGAG 8398
DB 9725 CATGCTATTAATCCCAACCTGTGGAGGCGAGGAGGCAATCACTTGAATCCAGAG 9666
QY 8399 TTTTATACAGCTGTGGCAACATGTGAATCTCATCTTCTAATAATAATAATAA--AT 8456
DB 9665 TTCATATATAGCTGTGGCAACATGTGAATCTCATCTTCTAATAATAATAATAA--AT 9606
QY 8457 TAGCAGAGTGTGGCATGCACTAGTATCCAGCTGTTCAAGAGGAGTGAAGGAGG 8516
DB 9605 CCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9546
QY 8517 ATTTGCTGAGTGTGGAGGCAAGATTTGCAATGAAACCGTGAATGTGCTTGTGAATCCAG 8576
DB 9545 ATGACTTGAACCGGAGGAGGTTACAGTGAAGCGAATTTGCACTGCACTCCAG 9486
QY 8577 CTTGGGTGACAGATTGAGACCTGTCTCAAAAAACAATTAATAATAATAATAAT 8633
DB 9485 CTTGGGTGACAGGAGCAAGATCTCATCTCAAAAAACAATTAATAATAATAATAAT 9429

Search completed: December 15, 2003, 12:37:36
Job time : 1612.46 secs



8006 --CCTCATAGCTTTTATCTATTCTCTTGCCTTCACACTTTGGCTTGCACGTGGAC 8063

Db	10845	GGACTCTGGTGGTTTGGCCAAA	CTTTCTTGAAGTGA	CACTAGACACTGAAATGTTCC	107866
Qy	8064	CATCATGTTCTTCCACTTTCT	CACACTACTGATCTTTCACTCAAGTTC	CACTACTGAT	8123
Db	10785	CTACCTGCTTCTTCTCTTCT	CTCTCTCTCCACAAGGGTTAACTTGACATGTC	-ACTTTC	107277
Qy	8124	ACCTCCCTCAGTTGCTCTTT	TTTCTTCTAGTAAAGATTTCCAGAGAGGGAAATCGAAATGGCCC	8183	
Db	10726	ATGGCTCTCCACAGCTCTCT	CAAGCCCTCTCCCATTTCTTTGCACCCAGCATTTTCTCTCA	10667	
Qy	8184	AGTCCATATTTTCAAGACA	CAACA	CATTTAAAGTGTGATTTGCCAGCCTATGTATTTGGCT	8243
Db	10666	ACACACCTCTTTGACAGCT	TAATCCACTTGTGTGCTGCTTCCAGAGGATCATTAAC	CTGA	106070
Qy	8244	ACAATTATGGGTGGGA	ACTCATCATTTACTTATCTTGCACAAAGCAGCATAGCTGTGTT	8303	
Db	10606	CACAACTAGGAAGAAGATCT	GAA--TACATGAGACATGCAATGATTAATGATGATGGGTA	105560	
Qy	8304	CTCAAAATAGAGCCCTCT	GGCCCAAGTGTGTGCTCATGCTTAAATCCCAACATCTGTG	8363	
Db	10549	CAAGTAAATAATGAAAGT	GGGCTGTGGCACGGTGGCTCACGTTGTATCCAGACATTTGG	10490	
Qy	8364	GAGGCCAGAGGGGGCAGAT	CACTTGAGTCCAGAGTTCTAGACACAGCCTGGGGCAACATGG	8423	
Db	10489	GAGGCCAGAGTGGGTGAT	CACCTGAGTCAAGAGTTTCAGACAGACCTGGCCAAATATGG	10430	
Qy	8424	TGAATATCATCTCTACTA	TAAAAATACAAAATTTAGCCAGGTGTGGTCATGACCACT	8483	
Db	10429	TGAATATCCATCTCTACTA	TAAAAATACAAAATTTAGCTGGGATGAGTGTGGCGGCACTGT	103707	
Qy	8484	AGTCCAGCTGTTCAAGAG	CGTGAAGTGGAGGATTTCTCGAATGTGGAGGCAAGATTT	8543	
Db	10369	AATGCCAGTACTTGGAG	CGTGAAGGAGGAGAAATTGCTTGAACCGAGAGGAGAGGTT	10310	
Qy	8544	GCAGTGAACCGTAC	TGCTGTGCAATCCACCTGGGTGACAAATTGAGACCTGTCT	8603	
Db	10309	GTAAGTGAATATATG	CCACCACTGTCACTCTAGCTGGGCGACAGAGGAGACTCATCT	10250	
Qy	8604	CAAAAAACAATTAATTA	ATAA	8626	
Db	10249	CAAAAAAGAAAAAAGT	GA	10227	
RESULT 2					
US-08-965-048-5					
Sequence 5, Application US/08965048					
Patent No. 6323244					
GENERAL INFORMATION:					
APPLICANT: Chem, Hong					
APPLICANT: Frelmet, Nelson					
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND					
FILE REFERENCE: 7853-093					
CURRENT APPLICATION NUMBER: US/08/965,048					
NUMBER OF SEQ ID NOS: 8					
SOFTWARE: Patentin Ver. 2.0					
SEQ ID NO 5					
LENGTH: 45716					
TYPE: DNA					
ORGANISM: Homo sapiens					
US-08-965-048-5					

Query Match	3.2%	Score 317.2;	DB 4	Length 45716;
Best Local Similarity	62.1%	Pred. No. 1.1e-64;		
Matches 614;	Conservative 0;	Mismatches 318;	Indels 56;	Gaps 5.

[illegible]

Db	20499	TGCAGTGATATGCATCTTGGCTCACTGCAACCTCCGCTCTGGGTTCAACAGTTCTCCT	20558
Qy	7772	GCCTCAGCCTCCGGTATAGCTGGAGCTTAAGAGCGCATACCACTCCGGCTAATTTT	7831
Db	20559	GCTTCAGCCTCCGGTATAGCTGGAGCTTAAGAGCGCATACCACTCCGAGC-----	20610
Qy	7832	GTAATTTTTTTTTGGTTACAATGTAATTAATTAATTTTGTATTTTGTATAGTA	7891
Db	20611	-----TAAAGTTTGTATTTTGTATAGTA	20632
Qy	7892	GATAGGGTTTCACCATGTTGGCCAGGCTGTCTCGAATCTCTGACCTCAGGTGATCCGC	7951
Db	20633	GATGGGGTTTGACCAAGTTGGCCAGGCTGTCTCGAATCTCTGACCTCAGGTGATCCGC	20692
Qy	7952	TGCTCGGCTCCCAAGTGTGAATTAACAGCATGAGCAACCAAGCTGGCC-CTTCA	8010
Db	20693	TTCTCAGCCTCCCAAGTGTGAATTAACAGTGTGAGCCACCAATGCCATCTTAGAGA	20752
Qy	8011	TAGTTTTTATCTATTCCTTTGGCTTCTCACAACTTTGGCTTGACGTGAGCAATCARG	8070
Db	20753	TGTTTTATATTTTAAAGTAAACAATTTTATTCAGTTAAATTCAGGCTTGAGTCAATTA	20812
Qy	8071	TTCTCTCACCCTTCTCACTACTTCATGATCTTTTCAGTCTCAGTTCCAACTGATACCTCC	8130
Db	20813	GATCATCAGATTTTGAAGTAAACATCTATTTCTGTAGACGTGATGATCTAAATGACCTC	20872
Qy	8131	TCAGTTGCTCTTTTTCTTAGTAGATTTCCAGAGAGGATCTGAATGGCCAGTCCAT	8190
Db	20873	AAGACTGAATTTAGCTGAATTAATAATTTGTCTAAGAGAGATAGGCTCTGAATTTGGGT	20932
Qy	8191	ATTTTCAGACCAACAACATTAAGTGTGATGTGCAGCCATGATTTGGTACATTA	8250
Db	20933	CTTAAGGCTATCTCTCTTAAGACCTCATCTCTCTCTTAATTTGAAATGCTC	20992
Qy	8251	TGGTTGGGAACTCATCATTT-----TACTTCATTCGCAAAAGACAGATAGCTGTGTTCC	8304
Db	20993	TAAAGATTAAGTACTTAATAGGCAATCTCTTATGAGAAATTAATGTAGATAGTGTAG	21052
Qy	8305	TCAAAATAGGCCCCCT-GGGCCAGGTGTGTGCTTATGCTTAATATCCCAACTGTGG	8363
Db	21053	AGATGTAAAGAAAGTCAAGGCGGGCGGCTGAGCTCACTCTGTATCCAGCACTTTGG	21112
Qy	8364	GAGGCCGAGGGGGGAGTCACTTGAATCAAGATTTTCAAGCCAGCTCGGGCAACARG	8423
Db	21113	GAGGCCGAGGGCAAGGATCAC-GAGGTCAAGAGGTCAAGACATCTGTGCTAACARG	21170
Qy	8424	TGAATATCTCATCTCTAATAAAAAATACAAAAATTAAGCCAGGTGTGTGTCATGCAACAGT	8483
Db	21171	TGAAACCCCGTCTCTATCTAAAAATACAAAAATTAAGCCGAGGATGTGTGCGGCGCCCAT	21230
Qy	8484	AGTCCCAAGCTGTTCAAGAGGCTGAGGTGGAGAGATGTCTGAGTGTGGAGGCAAGATT	8543
Db	21231	AGTCCCAAGTTACTCAAGAGGCTGAGGCAAGAGAAAGCGGTGAACCTGGGAGGCGAGCTT	21290
Qy	8544	GCAGTGAACCGTATCTGTGCTTCTGCAATCCAGCTTGGGTGACAAATTTGAGACCTGTCT	8603
Db	21291	GCAGTGAGCGAATGCAACACTGCACTCCAGCCTGGGCAACAAGTGTGAGCTGTCT	21350
Qy	8604	CAAAAAACAATTAATTAATAAATAAATAA 8631	
Db	21351	CAAAACAATAAAAAAAAAAAAAAAAAA 21378	

RESULT 3
US-08-965-048-6
; Sequence 6, Application US/08965048

```

: GENERAL INFORMATION:
:
: APPLICANT: Chen, Hong
: APPLICANT: Freilmer, Nelson
:
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
:
: TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
:
: FILE REFERENCE: 7853-093

```


[illegible]

```

: Parent No. 6323244
:
: GENERAL INFORMATION:
:
: APPLICANT: Chen, Hong
:
: APPLICANT: Freeman, Nelson
:
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
: TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
:
: FILE REFERENCE: 7853-093
:
: CURRENT APPLICATION NUMBER: US/08/965,048
:
: CURRENT FILING DATE: 1997-11-05
:
: NUMBER OF SEQ ID NOS: 8
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 5
:
: LENGTH: 45716
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: US-08-965-048-5

```

Query Match	3.1%;	Score 307.2;	DB 4;	Length 45716;
Best Local Similarity	62.1%;	Pred. No. 2.6e-62;		
Matches 616;	Conservative 0;	Mismatches 318;	Indels 58;	Gaps 6;

OY	7652	TTTTGTGTGTGTGTCGTGTGTGTGTAGANNGAGTCTACTCTGTCAACCAAGGCTGGAG	7711
Db	21377	TTTTTTTTTTTTTTTTTTTTTTTTTTTGTATGAGACAGAGTCTACTCTGTGTGCCAGGCTGGAG	2131
OY	7712	TGCATATGGCGCAATCTTGTGTTCACTGCAACCTCTGCTCTCGGTTCMAAGCATCTCTCT	7771
Db	21317	TGCAGTGTGTGAGATCTCGGCTCACTGCAACCTCGGCTCCCAAGTTCAACCGCTCTCTCT	2125
OY	7772	GCCCTACCCCTTCCGTAATAGCTGGGACTACAGGGCGATACCAACATGCTGTGGCTAATTTTT	7831
Db	21257	GCCCTACCCCTCTCGAAGTACCTGGGACTATATAGGGGCGCCGACACATGCTCGGCTAAT----	2120
OY	7832	GTATTTTTTTTTTGGGTTACAAATGTAATAATTTAATTTAATTTTGTATTTTATAGAGA	7891
Db	21201	-----TTTTGTATTTTATAGAGA	2118
OY	7892	GATAGGGTTTACCCANGTGGCCAGCGTGTCTGAACTCTGACCTCAAGTATCCGC	7951
Db	21182	GACGGGGTTTACCCATTTTATAGCAGAGATGGTCTGACCTCTGACCTC--GTATCTGCC	2112
OY	7952	TGCTCTCGGCTTCCCAAGTGTCTAGATTAACAGCATAAGCAACAGCGCTGGCC-----	8005
Db	21124	TGCTTGTGGCTTCCCAAGTGTCTGGGATTAACAGAGTAGGACACCGCGCTGGCTGTACCT	2106
OY	8006	CCTCATATGGTTTTATCTATTTCTTTTGTCTTTCACAACTTTTGGCTTGCAGCGTGACCA	8065
Db	21064	TCTCTTAATCTCTAACAATCTACATCTATTTTCTCATAAAGATTTGCCATTTTAGTA	2100
OY	8066	TCAATGTTCTCTCCAATTCTCTCACTACTATGATATTTTCACTTCACTCACTGTTCCAACTGATAC	8125
Db	21004	CTTATATCTCTCAGAG--GCACTTTTCCAAATTCAGGGAGAGAGATAGAGTCTCTTAGAGAGA	2094
OY	8126	CTCCCTCAGTTGTCTTTTTTTCCTAGTAATATTTCCAGAAAGGGAATCTGAATGGCCAG	8185
Db	20945	GATAGCTCTTAGAGCCCAAAATCTCAGAGCTCATCTCTTAGTAGAACTATTTATTTAGC	2088
OY	8186	TCCATATTTTCAGACACACCACTTAAATGTGTGATTTCCAGCCATGATATTTGGCTAC	8245
Db	20885	TAAATTCAGTCTTAGAGCATTTTAGATATCAAGCTTTACAAATTAATGAGTTTTCACCTCAA	2082
OY	8246	ATTAAATGGGTTGGAACTCATATTTTACTTATTTGCAAAAGCAGCATAGCTCTGGTCT	8305
Db	20825	AATATCATGATCTAAATATGACTCAAGCCTGAATTTTAACTGAATTAATGTTTTACTTTAA	2076
OY	8306	CAAAATATGGGCC---CTGGGCCAGAGTGTGTGTGCTCATGCTATATATCCCAACAATGTG	8362
Db	20765	AATATTAACAACAATCTCTAGAGATGGGATGAGTGTGCTCAACCTGTAAATCCAGACATTTG	2070
OY	8363	GGAGGCGAGAGGGGGAGATATCATTTGAGTCCAGAGATTTCTAGAACCAAGCTGGGCAACATG	8422
Db	20705	GGAGGCGTGGAGAGGCGGATATCATTTAGAGTCCAGAGTTTCAGAACCAAGCTGGGCAACATG	2064

QY 2401 ACCTCAACCCAGACGCTGCTAGATGCTGAGCTTTGATCTACCACTTGATGAC 2460
Db 51127 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 51068
QY 2461 AATGGAACCATGCTGTTT 2478
Db 51067 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 51050

RESULT 5
AX345846
LOCUS AX345846 5234 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 917 from Patent WO0200928.
ACCESSION AX345846
VERSION AX345846.1 GI:18493732
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 917 03-JAN-2002;
Epidemiol. Infect. 124: 1-10 (2000)
FEATURES
source 1. 5234
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1479 a 40 c 1361 g 2354 t
ORIGIN

Query Match 15.1%; Score 1509.4; DB 6; Length 5234;
Best Local Similarity 79.9%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 446; Indels 0; Gaps 0;

QY 12 AGGCAAGCCAGCAAAATCTGTAGATTCAGACGGGCTTCAAGCTGACCAAGTTG 71
Db 3012 AGGTAGATTAGTAATTTGTAGATTTGATGAGTTGATGAGTTGATGAGTTGATGAGTTG 3071
QY 72 TTGAGAAATTTCTGTAGAGATCATGGGCTCTCAGAGGAGAAATATAGGTTTCA 131
Db 3072 TTGAGAAATTTCTGTAGAGATCATGGGCTCTCAGAGGAGAAATATAGGTTTCA 3131
QY 132 AGCTAGAGGAGAAAGAGTGAAGGAGGAGCTTGAATAGTGGTCCCATTTGCCCA 191
Db 3132 AGCTAGAGGAGAAAGAGTGAAGGAGGAGCTTGAATAGTGGTCCCATTTGCCCA 3191
QY 192 CACCCAGAAAGAGACATGCTGCAATGGGAGAGAGTATGATGACATTTGGCTGA 251
Db 3192 TATTTAGAAAGAGATATGTTTGTATGAGGAGAGAGTATGATGATATGTTGTA 3251
QY 252 GCAGCAGTGCATTCGCCAGGCTGCCAAGACTCAGAGAGTCCAGCTTGCCTGACC 311
Db 3252 GTAGCAGTGCATTCGCCAGGCTGCCAAGACTCAGAGAGTCCAGCTTGCCTGACC 3311
QY 312 TATGAGAGGAGATGATGTTCAAGCATTTCCTGTAAGTCAAGAGAGAGATTTGA 371
Db 3312 TATGAGAGGAGATGATGTTCAAGCATTTCCTGTAAGTCAAGAGAGATTTGA 3371
QY 372 GCTGTAGGAGAGGCTGTGATCATGTTTCCAGAGTTCCGGAATGTTGTTTCT 431
Db 3372 GTTGTAGTGAAGGTTTGTGATGATGTTTTCAGAGTTTCCGGAATGTTGTTTCT 3431
QY 432 GTTGAAGAGAACTTGCAGAGTGAAGAGAGCTGAGACTTTTGTGTAATATATAT 491
Db 3432 GTTGAAGAGAACTTGCAGAGTGAAGAGAGCTGAGACTTTTGTGTAATATATAT 3491
QY 492 GGGAGTGTGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 551
Db 3492 GGGAGTGTGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 3551

QY 552 GGAGCAGGGGCTCTGACATTCATCTGCGCCATAGCTAGTCTGCCATATTTCTTC 611
Db 3552 GGAGTGGGGTTTGGAAATTTATTTGTTATATGTTAGTTTGTATATATTTTTC 3611
QY 612 TGACTCACTAGGCAATCTCAACAGAAATGGGGAGCTTTGGAGTGGGCCAGAGAT 671
Db 3612 TGAATTTATAGGTAATTTATATAGAAATGGGGGTGTTTGGAGTGGGGTTTGAAG 3671
QY 672 ACTGAGATAGCAGGATGATCCAGAGAGATGAGATGGGGCCGAGACCTGAGAG 731
Db 3672 ATTGAGATAGTGAAGATTTTGAAGAGATGAGATGGGGCTGAGATTTGAGAG 3731
QY 732 AGAAGAGAGCTGATGATGAAGAGGCTGCTGACTCTGATCTGATTTAGCTTAC 791
Db 3732 AGAAGATAGATTTGATGATGAAGGCTTGTGATTTTATGATTTATATATAT 3791
QY 792 CATACCTTACCCCAATTCCTAATTTCTCTAACCCTAGAGGGGGGAAATGTCAGAA 851
Db 3792 TATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3851
QY 852 TTTGGCTGACACTGACCAACTACTGACTGATGAATGATTTTGCATTTTTC 911
Db 3852 TTTGGTTGATATGATGATATTTATTTATTTATTTATTTATTTATTTATTTATTT 3911
QY 912 TTCAACAATATTTCTGAAACAATCTTATATGCAAGCACTTTTATGAGTCAAGGAT 971
Db 3912 TTTTAAATATTTTATTTGAAATATTTTATATTTATTTATTTATTTATTTATTTATTT 3971
QY 972 ATATATGTTAAACAAGACAGGCAAAACAAGCAACAACAATCACCAGATTA 1031
Db 3972 ATATATGTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4031
QY 1032 GTAGACAGATGAAGAAATTTCAAGTTTATGATGATGATGATGATGATGATGATGAT 1091
Db 4032 GTAGATGATGAAGAAATTTTATGTTTATGATGATGATGATGATGATGATGATGAT 4091
QY 1092 ATGCTAGATGAAGCCGCTCAAGAAAGCTTCAATGAGAGTATGATTTAAGCAGAGTC 1151
Db 4092 ATGCTAGATGAAGCCGCTTAAAGAGTTTATGAGAGTATGATTTAAGAGAGTT 4151
QY 1152 AGCTAGAAATTTGTAATTTCAAGTTCAATTTGTTGCTGGTGTAAATTA 1211
Db 4152 AGTTAGAAATTTGTAATTTTATGATGATGATGATGATGATGATGATGATGATGAT 4211
QY 1212 GCTTTTCCCAAGGTGAACTACAGAAAGATTTATGATGATGATGATGATGATGAT 1271
Db 4212 GTTTTATTTTATGAGTGAATTTATTAAGAAATTTATTTATGATGATGATGATGAT 4271
QY 1272 TCTGAGAGAGACACTCTGTTCTGCTCATTTACTGTAACCTTCCAGGCA 1331
Db 4272 TTTGAGAGAGATATTTTGTGTTTGTATTTATTTATTTATTTATTTATTTATTT 4331
QY 1332 CTTTTCAGAGCCCTTTCAGCTGAGAGAGGAGGCTGAGGCTTGGACA 1391
Db 4332 TTTTGTGTAAGTTTGTGTTGATGAGAGGCTGAGGCTTGGAGAT 4391
QY 1392 TTTGAGAGAGATATTTTGTGCTCCAGACTCACTGAGCCCAAGGTTAAAGTAA 1451
Db 4392 TTTGATTAATGATGATATTTTGTGTTTATTTATTTATTTATTTATTTATTTATTT 4451
QY 1452 GGTGTGAGATGCTCCAGAAAGGCTCTGATGAGGCTTGAAGAACTGTTCTGAT 1511
Db 4452 GGTGTGAGATGCTTTCAGAAAGGCTTTCATGAGGCTTGAAGAAAGTGTGTTTGA 4511
QY 1512 TTTCTAAGTATTTATCTGTGTGTGTTTCTAAGATGATGATGATGATGATGATGAT 1571
Db 4512 TTTTAAAGTATTTATTTGTGTGTGTTTCTAAGATGATGATGATGATGATGATGAT 4571
QY 1572 GCTGCTTGAAGAGGCTTGAACAAGTCAACAAGCAATTTAGTCCCTTTCTCAC 1631
Db 4572 GTTGTTTAGAGAGGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 4631
QY 1632 CTCCTCATTTGAGATTAATCAATCAAAAGGTGATCTCACTGATCTTCCC 1691

Db 53347 CCATTGCCCAACACCCAGAAAGACATGCTCCGATGAGGAGAGGATGATAGCA 53288
Qy 241 CATTGGCTGATGAGGATGGCATTTGCCAGCTGCAAGAGACTCAGAGGTCACGCTT 300
Db 53287 CATTGGCTGATGAGGATGGCATTTGCCAGCTGCAAGAGACTCAGAGGTCACGCTT 53228
Qy 301 GCCCACTGACCTATGAGAGGAGATGATTTTCAACAGACATTTTCAATTCGTAAAGTACGA 360
Db 53227 GCCCACTGACCTATGAGAGGAGATGATTTTCAACAGACATTTTCAATTCGTAAAGTACGA 53168
Qy 361 GAGGACATTTGAGCTGATGAGGAGGCTGAGTCAATTTTTCAGAGGTTCCGGAATG 420
Db 53167 GAGGACATTTGAGCTGATGAGGAGGCTGAGTCAATTTTTCAGAGGTTCCGGAATG 53108
Qy 421 TGTGTTTCTGTTGAGAGGAACTTCGAGATGAGAAAGGATCGAGACTTTTGTGTA 480
Db 53107 TGTGTTTCTGTTGAGAGGAACTTCGAGATGAGAAAGGATCGAGACTTTTGTGTA 53048
Qy 481 AGATTATATATGAGGATCTGACAGGAGCTGAGGCACTGTGAGGAGTCAAGGCTTTTCA 540
Db 53047 AGATTATATATGAGGATCTGACAGGAGCTGAGGCACTGTGAGGAGTCAAGGCTTTTCA 52988
Qy 541 GCCTTGGCTAGGAGGAGGAGGCTCTGAACTTCTGAGGCTGAGTCTGAGTCTGAGT 600
Db 52987 GCCTTGGCTAGGAGGAGGAGGCTCTGAACTTCTGAGGCTGAGTCTGAGTCTGAGT 52928
Qy 601 TAATTTCTTTTCTGACCTCAGAGCAATCTCAGAGAAATGGGGAGCTTTGGAGTGG 660
Db 52927 TAATTTCTTTTCTGACCTCAGAGCAATCTCAGAGAAATGGGGAGCTTTGGAGTGG 52868
Qy 661 GCCCAGAGAGTACTGAGATGAGAGGATGAGATCCAGAGAAAGATGAGTGGGGCCGAG 720
Db 52867 GCCCAGAGAGTACTGAGATGAGAGGATGAGATCCAGAGAAAGATGAGTGGGGCCGAG 52808
Qy 721 ACACTGAGAGAGAAACAGGACTGTGAGATTAAGGCGTGTGACTCTGAGTCTGAT 780
Db 52807 ACACTGAGAGAGAAACAGGACTGTGAGATTAAGGCGTGTGACTCTGAGTCTGAT 52748
Qy 781 ATGCCCTACTACCACTAATCCCAATTCCTTAATTTCTCCCTAAGGAGGGGGGAA 840
Db 52747 ATGCCCTACTACCACTAATCCCAATTCCTTAATTTCTCCCTAAGGAGGGGGGAA 52688
Qy 841 ATTTGTCAGAAATTTGGCTGACACTAGCAACACTACTAGTACTGAAATGCAATTTTGG 900
Db 52687 ATTTGTCAGAAATTTGGCTGACACTAGCAACACTACTAGTACTGAAATGCAATTTTGG 52628
Qy 901 CATTTTTTCATTCAACAATATTTTCTGAAACAATCTTATATGCAAGGCACTAATTTTGA 960
Db 52627 CATTTTTTCATTCAACAATATTTTCTGAAACAATCTTATATGCAAGGCACTAATTTTGA 52568
Qy 961 GAGTCAGGATATATATATGTTAAACAAGCAGGCAAAACAAGCAAGCAACAACA 1020
Db 52567 GAGTCAGGATATATATATGTTAAACAAGCAGGCAAAACAAGCAAGCAACAACA 52508
Qy 1021 TCACAGATATATGAGACAGATGAAAGAAATTTCAAGTTTATGTAATTAATAACAAG 1080
Db 52507 TCACAGATATATGAGACAGATGAAAGAAATTTCAAGTTTATGTAATTAATAACAAG 52448
Qy 1081 AAGGGCTGTAATATGCTAGATTAAGGCGCTCAGAAAGGCTTCAATGAGAGGTAGCATTT 1140
Db 52447 AAGGGCTGTAATATGCTAGATTAAGGCGCTCAGAAAGGCTTCAATGAGAGGTAGCATTT 52388
Qy 1141 AAGCAGAGATCAGCTAGAAATATTTGAAATTCAGTTACAGTTTATTTTGTGGGTT 1200
Db 52387 AAGCAGAGATCAGCTAGAAATATTTGAAATTCAGTTACAGTTTATTTTGTGGGTT 52328
Qy 1201 GGTAAATATATAGGTTTCCCAAGGTGAAATCTCAAGAAAGCTAATTAATCTAGTAG 1260
Db 52327 GGTAAATATATAGGTTTCCCAAGGTGAAATCTCAAGAAAGCTAATTAATCTAGTAG 52268
Qy 1261 TGTGTGTGTCTCTGAGAGAGACACTCTGTGTCTGTCTGATTAATCTGCAACCTTTC 1320
Db 52267 TGTGTGTGTCTCTGAGAGAGACACTCTGTGTCTGTCTGATTAATCTGCAACCTTTC 52208

Qy 1321 ACTTCAGGCACTTTTGGAAAGCCCTTGGCCAGTCAAGGAAAGCCAGAGGCTGGGCAATG 1380
Db 52207 ACTTCAGGCACTTTTGGAAAGCCCTTGGCCAGTCAAGGAAAGCCAGAGGCTGGGCAATG 52148
Qy 1381 GGGCTTGGACATTTTGAACAACAGTGAACATTAATTTTCCCAAGCTCAGTACCCCAAGGTT 1440
Db 52147 GGGCTTGGACATTTTGAACAACAGTGAACATTAATTTTCCCAAGCTCAGTACCCCAAGGTT 52088
Qy 1441 AAAGCTGAAGAGCTTGGGCAATGCCAGAAAGCCCTGATGAAGCTTGAAAGAAAGCTG 1500
Db 52087 AAAGCTGAAGAGCTTGGGCAATGCCAGAAAGCCCTGATGAAGCTTGAAAGAAAGCTG 52028
Qy 1501 TTCTGAGATATTTCTAATTAAGTTATCTGATGAGTGGTACTCTAATTAAGTAAAGTAT 1560
Db 52027 TTCTGAGATATTTCTAATTAAGTTATCTGATGAGTGGTACTCTAATTAAGTAAAGTAT 51968
Qy 1561 TGTGTCTGTAGCTGCTTGAAGCAGAGGCTTGAACAAGTACACAGCAATATTAGTCCCT 1620
Db 51967 TGTGTCTGTAGCTGCTTGAAGCAGAGGCTTGAACAAGTACACAGCAATATTAGTCCCT 51908
Qy 1621 CCTTTTCTCACTCCGCCATTTGTGAGATTAACATCAACAAGGTATCTCAGTCT 1680
Db 51907 CCTTTTCTCACTCCGCCATTTGTGAGATTAACATCAACAAGGTATCTCAGTCT 51848
Qy 1681 ACTCACTTCCCTGACTTATGAGATGCTGAGACCAATGACAGTGTAGAGTCAAGCTGGA 1740
Db 51847 ACTCACTTCCCTGACTTATGAGATGCTGAGACCAATGACAGTGTAGAGTCAAGCTGGA 51788
Qy 1741 CGTCAAGCACTGATGCCCACTTACTGTTGAAATTCCTGAAGGGGGTGGGGGAGCTG 1800
Db 51787 CGTCAAGCACTGATGCCCACTTACTGTTGAAATTCCTGAAGGGGGTGGGGGAGCTG 51728
Qy 1801 CCGGAAAAAGAGATCTTGGATTCAGATTTCTGTCCAGACCTGACCTTATTTTGCAGTGA 1860
Db 51727 CCGGAAAAAGAGATCTTGGATTCAGATTTCTGTCCAGACCTGACCTTATTTTGCAGTGA 51668
Qy 1861 TGTAAATCAGCAATATTTGGCTTATGCTGGGAGACAGACATTTCCAGTAAAGTGGAGG 1920
Db 51667 TGTAAATCAGCAATATTTGGCTTATGCTGGGAGACAGACATTTCCAGTAAAGTGGAGG 51608
Qy 1921 TGGGGGTGTGTGTCTGCCAATCTATATAGGAGTTCACTGTCTCAGACAGCTGTCTC 1980
Db 51607 TGGGGGTGTGTGTCTGCCAATCTATATAGGAGTTCACTGTCTCAGACAGCTGTCTC 51548
Qy 1981 TGTGGCTCTGACGCTCAGACATGCGTGAAGGTACTGAGACACCCGTTGACTGGGGTTGT 2040
Db 51547 TGTGGCTCTGACGCTCAGACATGCGTGAAGGTACTGAGACACCCGTTGACTGGGGTTGT 51488
Qy 2041 GGAAGCTATGCTGTCTGAGCAATGCAACCCCTTCTCCGATGAGTAAAGTGGAGC 2100
Db 51487 GGAAGCTATGCTGTCTGAGCAATGCAACCCCTTCTCCGATGAGTAAAGTGGAGC 51428
Qy 2101 TAGAAGCAGAGATTAAGTTCTGAGGCTAGGAGTAAAGGAGGCAATTTTGAAGGCTCGGT 2160
Db 51427 TAGAAGCAGAGATTAAGTTCTGAGGCTAGGAGTAAAGGAGGCAATTTTGAAGGCTCGGT 51368
Qy 2161 CAAATTTGGGGTCAAGGGCTATGAGAAAGGATCGGTCCCAATGATCAAGATATCTAAT 2220
Db 51367 CAAATTTGGGGTCAAGGGCTATGAGAAAGGATCGGTCCCAATGATCAAGATATCTAAT 51308
Qy 2221 TTGTTCTCCCTGAGACTGATGCGCATGAGGAATGTTGGAAGGCAAGCAAGCAAGCAGC 2280
Db 51307 TTGTTCTCCCTGAGACTGATGCGCATGAGGAATGTTGGAAGGCAAGCAAGCAAGCAGC 51248
Qy 2281 CAGAGCTGACTGTAGAGGCTGACTGCTAGTCTGTCTATCTGTCTGTGTGTCTC 2340
Db 51247 CAGAGCTGACTGTAGAGGCTGACTGCTAGTCTGTCTATCTGTCTGTGTGTCTC 51188
Qy 2341 TGCATTTTATCACTTGTGGTTTTTTTTTTTTTTTTTTTTTTTTTACTTGCATCTCCCT 2400
Db 51187 TGCATTTTATCACTTGTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 51128

```

Db      19624 AATGTGGACATGCTGTATCTCAGCTTACTTGGAGGCTGAGGAGAGATCGCTT 19683
Qy      9677 GAACCTGGAGGATGATGTTGCGATGAGCCGAGATTGACCATTCCTCAGCTCGGG 9736
Db      19684 GAACCTGGAGGATGATGTTGCGATGAGCCGAGATTGACCATTCCTCAGCTCGGG 19743
Qy      9737 AACAGAGTGAATCTCCGTCTTAATTAATGAATGAATGAATGATATCAGTCAAGTAG 9796
Db      19744 AACAGAGTGAATCTCCGTCTTAATTAATGAATGAATGAATGATATCAGTCAAGTAG 19803
Qy      9797 GGAAGGGAAGAGGCTTCAAGATGACTGAGCTTTGCTGACTCAGCAACTGAGTGCT 9856
Db      19804 GGAAGGGAAGAGGCTTCAAGATGACTGAGCTTTGCTGACTCAGCAACTGAGTGCT 19863
Qy      9857 GGTGCTTTGTTTCTTAAATTTGGAAAGCTAGGAGTGTGTGTGTGTGTGGGGGGCAG 9916
Db      19864 GGTGCTTTGTTTCTTAAATTTGGAAAGCTAGGAGTGTGTGTGTGTGTGGGGGGCAG 19923
Qy      9917 AATCAGTTGGGATATTAGTTTGGGTGCTTATGCGACCCCATTAAGCATGTCAAGT 9976
Db      19924 AATCAGTTGGGATATTAGTTTGGGTGCTTATGCGACCCCATTAAGCATGTCAAGT 19983
Qy      9977 AGGCAGCTGATTGTGAGCTTAAC 10000
Db      19984 AGGCAGCTGATTGTGAGCTTAAC 20007

```

```

RESULT 4
AC024232/c 148198 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 11 clone RP11-223K12, WORKING DRAFT
DEFINITION
ACCESSION AC024232.3 GI:8516169
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 148198)
AUTHORS Waterson,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 148198)
AUTHORS Waterson,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2000) Genome Sequencing Center, Washington
UNIVERSITY School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jun 14, 2000 this sequence version replaced gi:7715652.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH022K12
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.99019
Consensus quality: 142463 bases at least Q40
Consensus quality: 144172 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 147398; sum-of-contigs
Quality coverage: 4.53 in Q20 bases; agarose-fp
Quality coverage: 4.67 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces

```

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2755: contig of 2755 bp in length
2855: gap of unknown length
2856: contig of 9398 bp in length
12254: gap of unknown length
12254: contig of 10214 bp in length
22568: gap of unknown length
22568: contig of 14935 bp in length
37603: gap of unknown length
37603: contig of 13351 bp in length
51054: gap of unknown length
51154: contig of 15540 bp in length
66694: gap of unknown length
66794: contig of 19087 bp in length
85880: gap of unknown length
85881: contig of 29975 bp in length
115956: gap of unknown length
116056: contig of 32143 bp in length.
Location/Qualifiers
1. 148198
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-223K12"
1. 2755
/note="assembly_name:Contig4"
2856. 12253
/note="assembly_name:Contig5"
12354. 22567
/note="assembly_name:Contig6"
22668. 37602
/note="assembly_name:Contig7"
37703. 51053
/note="assembly_name:Contig8"
51154. 66693
/note="assembly_name:Contig9
vector_side:right"
66794. 85880
/note="assembly_name:Contig10"
85981. 115955
/note="assembly_name:Contig11
clone_end:SP6
vector_side:right"
116056. 148198
/note="assembly_name:Contig12"

```

```

BASE COUNT 38577 a 34635 c 33776 g 40404 t 806 others
ORIGIN

```

```

Query Match 23.7%; Score 2374.8; DB 2; Length 148198;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2376; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
Qy      1 TCCCTTCCCGACGAGCCGACCAAAATCTGAGATTTCAGACAGGTTCTGACAGCTG 60
Db      53527 TCCCTTCCCGACGAGCCGACCAAAATCTGAGATTTCAGACAGGTTCTGACAGCTG 53468
Qy      61 AAGACAGTTGTTGAGAAATTCCTGATGAGATCATGGGCTGCTCAGAGGAGAAATA 120
Db      53467 AAGACAGTTGTTGAGAAATTCCTGATGAGATCATGGGCTGCTCAGAGGAGAAATA 53408
Qy      121 TAGGTTTCAAGAGCTGAGAGGAGAAAGAAAGGTGAGGGGAGCTTAAATATGTGCTC 180
Db      53407 TAGGTTTCAAGAGCTGAGAGGAGAAAGAAAGGTGAGGGGAGCTTAAATATGTGCTC 53348
Qy      181 CCAATGCCCAACACCCAGAAAGACATGCCCTGCAATGGGAGAGGTGACTATGAGA 240

```

[illegible]

Db	18546	AGAGATTGGCACTGAACCGTGA	CTGTGCTCAAA	CAATTAATTA	AAATTAATTAATATG	GGTTCCTGAGCA	GCGTAA	TTT	1860					
QY	8597	CTGTGCTCAAA	CAATTAATTA	AAATTAATTAATATG	GGTTCCTGAGCA	GCGTAA	TTT	8656						
Db	18606	CCTGCTCAAAA	CAATTAATTA	ATTAATTAATTAATATG	GGTTCCTGAGCA	GCGTAA	TTT	1866						
QY	8657	CAGTGGGAAC	CTTCCCA	GCGGGAAGGTGGA	TATGTCA	GTCAC	CCGCTGTAT	TACTCAGTACACG	8716					
Db	18666	CAGTGGGAAC	CTTCCCA	GCGGGAAGGTGGA	TATGTCA	GTCAC	CCGCTGTAT	TACTCAGTACACG	1872					
QY	8717	GCTAATAAGAA	CTTGTG	GAGCAGCA	GAAGAACACTGTG	ATTTACTCA	CAAAATAT	TTTG	8776					
Db	18726	GCTAATAAGAA	CTTGTG	GAGCAGCA	GAAGAACACTGTG	ATTTACTCA	CAAAATAT	TTTG	1878					
QY	8777	TTGAGCAT	CTGATTA	GAGAGTGGG	CACTTGTCTTA	GGCACTGAGAT	ACAGTACA	ATCG	8836					
Db	18786	TTGAGCAT	CTGATTA	GAGAGTGGG	CACTTGTCTTA	GGCACTGAGAT	ACAGTACA	ATCG	18844					
QY	8837	CAGACAA	AGATGCC	TGCTCCCTG	ACAGGCTCTG	CTAAGTGA	AGAGACATTA	AGAAGAGA	8896					
Db	18846	CAGACAA	AGATGCC	TGCTCCCTG	ACAGGCTCTG	CTAAGTGA	AGAGACATTA	AGAAGAGA	1890					
QY	8897	AAGGA	AAAAAGAGA	TAA	TTTTTGGTAA	TTAAGGTTGTAA	AGAAAAATTA	AGACAGAT	8956					
Db	18906	AAGGA	AAAAAGAGA	TAA	TTTTTGGTAA	TTAAGGTTGTAA	AGAAAAATTA	AGACAGAT	18966					
QY	8957	AGTGGATTA	GAGGTGAGAGAA	TGAGGG	CTGTCTTCTGA	AGAAATGA	TTTTTGGCTGAG		9016					
Db	18966	AGTGGATTA	GAGGTGAGAGAA	TGAGGG	CTGTCTTCTGA	AGAAATGA	TTTTTGGCTGAG		1902					
QY	9017	ACTTCAGT	GTATGAGAGAA	TTTAA	CCACAGATGTCTG	GAGAGAAA	GCATTTT	AGGAG	9076					
Db	19026	ACTTCAGT	GTATGAGAGAA	TTTAA	CCACAGATGTCTG	GAGAGAAA	GCATTTT	AGGAG	1908					
QY	9077	GGTGA	GAGACA	CA	TACTTCAAGGA	ATCAAGAGAA	CCGTGTGA	GCGCTGAA	CA	CA	CAG	9136		
Db	19086	GGTGA	GAGACA	CA	TACTTCAAGGA	ATCAAGAGAA	CCGTGTGA	GCGCTGAA	CA	CA	CAG	19144		
QY	9137	AAAGAG	CAGGTGGGTG	ACTTGA	AAAGG	CAAGGAC	CGCAGTGG	CAGGTTA	CCT	TGA	CTG	9196		
Db	19146	AAAGAG	CAGGTGGGTG	ACTTGA	AAAGG	CAAGGAC	CGCAGTGG	CAGGTTA	CCT	TGA	CTG	1920		
QY	9197	GTAAGG	GT	TTTCA	AAACATTA	AAAGGAGT	CA	TCA	GA	AAAGTCTT	TGA	CA	GCGCTGTGATTA	9256
Db	19206	GTAAGG	GT	TTTCA	AAACATTA	AAAGGAGT	CA	TCA	GA	AAAGTCTT	TGA	CA	GCGCTGTGATTA	1926
QY	9257	TTTCTA	CTCAT	TTTTTTTA	TAAAGTCA	CTGCACTTTT	GCAGAA	CA	TTAGTTA	TAAAG		9316		
Db	19266	TTTCTA	CTCAT	TTTTTTTA	TAAAGTCA	CTGCACTTTT	GCAGAA	CA	TTAGTTA	TAAAG		1932		
QY	9317	TACAAG	CATATTA	AGCAAGAA	TCCAGCTAG	CAATCCGTCAG	TGTGCCAA	TT	AGAGSTG		9376			
Db	19326	TACAAG	CATATTA	AGCAAGAA	TCCAGCTAG	CAATCCGTCAG	TGTGCCAA	TT	AGAGSTG		1938			
QY	9377	ATGA	CCGCTTGA	CTAGAGT	ATAGCAGCAGAG	TGGTGAAGAT	CA	CA	TGA	TATATTT	9436			
Db	19386	ATGA	CCGCTTGA	CTAGAGT	ATAGCAGCAGAG	TGGTGAAGAT	CA	CA	TGA	TATATTT	19444			
QY	9437	TGAGG	GTAGACTG	ACATTA	ATAGCTTA	ATAGTGG	CCGGGTGTG	GTG	CTTA	CG	9496			
Db	19446	TGAGG	GTAGACTG	ACATTA	ATAGCTTA	ATAGTGG	CCGGGTGTG	GTG	CTTA	CG	19500			
QY	9497	CCGTAT	TCTTAG	CACTTTGG	AGGCGCAAGG	CGAGTGAAT	CA	CTG	AGGTC	CAGAGTTG	9556			
Db	19504	CCGTAT	TCTTAG	CACTTTGG	AGGCGCGCAAGG	CGAGTGAAT	CA	CTG	AGGTC	CAGAGTTG	19566			
QY	9557	AGAC	CAGCTTGA	CCAA	CATGTG	TAAAC	CTTGCTCTTA	CTTA	AAAAATTA	CAAAATTA	AGCTGG	9616		
Db	19564	AGAC	CAGCTTGA	CCAA	CATGTG	TAAAC	CTTGCTCTTA	CTTA	AAAAATTA	CAAAATTA	AGCTGG	19622		
QY	9617	AATGTG	GCACAT	GTCTGTAT	CTC	CAGCTTA	CTTTGG	AGGCTG	AGGCA	AGAGAT	GTGCTT	9676		

Db 15251 GAGCAAGCATATGAGAGGGGTAACCACTCTCAACCAATGCTGTGATGATTGAGACA 15310
QY 5297 AAGGTACCGCGAGTGGTCTTGAGGACTTGAGGCTACGGGAAACATGAAGAGAGCTCCGGC 5356
Db 15311 AAGGTACCGCGAGTGGTCTTGAGGACTTGAGGCTACGGGAAACATGAAGAGAGCTCCGGC 15370
QY 5357 CAGCTGTGGGAACTGTCTCTCTGCTGAGATGAGCTGGGCGCTACTAGCTTCCAGG 5416
Db 15371 CAGCTGTGGG-ACTGTCTCTCTGCTGAGATGAGCTGGGCGCTACTAGCTTCCAGG 15429
QY 5417 GTTACCAATTCCTGCGCTTGCACCTCTGTCAAGGAGAGGTGCTCCAGGTACCCGGGG 5476
Db 15430 GTTACCAATTCCTGCGCTTGCACCTCTGTCAAGGAGAGGTGCTCCAGGTACCCGGGG 15489
QY 5477 ATGTCCGAGACTTCAATGCGCTGCGCTGGCAGAGGTGAGAAAGCCCTAGCACTTGAGA 5536
Db 15490 ATGTCCGAGACTTCAATGCGCTGCGCTGGCAGAGGTGAGAAAGCCCTAGCACTTGAGA 15549
QY 5537 CCTGTGAGAAATTCATCCACTTTCCTGAGCTGTGTGATCTCAAGTGTCTAGCTTCACT 5596
Db 15550 CCTGTGAGAAATTCATCCACTTTCCTGAGCTGTGTGATCTCAAGTGTCTAGCTTCACT 15609
QY 5597 TTAACTCCGTGTGCGACACCTTGCGCTTAAATCTAGCCCAATTCCTAGATTTTC 5656
Db 15610 TTAACTCCGTGTGCGACACCTTGCGCTTAAATCTAGCCCAATTCCTAGATTTTC 15669
QY 5657 CCATTGCGCTCATATGGGGGAAACCAACCCCACTAACCCAGCGATCTTCCACTTG 5716
Db 15670 CCATTGCGCTCATATGGGGGAAACCAACCCCACTAACCCAGCGATCTTCCACTTG 15729
QY 5717 GACCTGACTGACCTCTGCGCTCTGTGTGTTCTCTGACCCATTTCTCTCCAGGC 5776
Db 15730 GACCTGACTGACCTCTGCGCTCTGTGTGTTCTCTGACCCATTTCTCTCCAGGC 15789
QY 5777 CATGACACAGGAATGGAGTGGCCATGGGAAAGTACCAACATGGCCCTGAGTATAG 5836
Db 15790 CATGACACAGGAATGGAGTGGCCATGGGAAAGTACCAACATGGCCCTGAGTATAG 15849
QY 5837 CGCTGAGCCCATCATGCTGTGTGCTGTGACATGACGTGTGACCAATGGTGGCACCTAT 5896
Db 15850 CGCTGAGCCCATCATGCTGTGTGCTGTGACATGACGTGTGACCAATGGTGGCACCTAT 15909
QY 5897 GCGTTGAGTGTGAGAGATGCGCCCAACTGCCCAATGTGCTGTCACTCTTTTAACTT 5956
Db 15910 GCGTTGAGTGTGAGAGATGCGCCCAACTGCCCAATGTGCTGTCACTCTTTTAACTT 15969
QY 5957 GTATCTCCCATCTTGTGACATTTCTGCATGTGTATCATCTGTGCTCACTTATTTGCC 6016
Db 15970 GTATCTCCCATCTTGTGACATTTCTGCATGTGTATCATCTGTGCTCACTTATTTGCC 16029
QY 6017 TGTGTCCCATCTTGTGACATGCGCTTGTGATCCCATCTGTGAGGCAATTTTCCAA 6076
Db 16030 TGTGTCCCATCTTGTGACATGCGCTTGTGATCCCATCTGTGAGGCAATTTTCCAA 16089
QY 6077 TCTTGTCTGTGACGCGCCAGCCCTTAATCTTCACTTACCTGTACCATTTCCCAAG 6136
Db 16090 TCTTGTCTGTGACGCGCCAGCCCTTAATCTTCACTTACCTGTATCTTCCCAAG 16149
QY 6137 GCTGTGCGCTGTGTGACCTCTGTGGGCGCTTATGACTCTTGTGTCTCTTGTCTAAT 6196
Db 16150 GCTGTGCGCTGTGTGACCTCTGTGGGCGCTTATGACTCTTGTGTCTCTTGTCTAAT 16209
QY 6197 GCGCTGTGAGCGCTGTGAGCTCTCCGTGTGCTCCCTGAGACTTATGTGTCTGTACCTC 6256
Db 16210 GCGCTGTGAGCGCTGTGAGCTCTCCGTGTGCTCCCTGAGACTTATGTGTCTGTACCTC 16269
QY 6257 CTGTGCTCCCTTGTCTGTGACATTTTCTGAGTCTCTGTGCTCCCTGATTTATCTT 6316
Db 16270 CTGTGCTCCCTTGTCTGTGACATTTTCTGAGTCTCTGTGCTCCCTGATTTATCTT 16329
QY 6317 CAGAACTCCATCTTGTGTTTCAAGTGTCTGTGTTCTTATGTCCAGACCCCTGGGCAATAGACT 6376
Db 16330 CAGAACTCCATCTTGTGTTTCAAGTGTCTGTGTTCTTATGTCCAGACCCCTGGGCAATAGACT 16389

QY 6377 GCCTGGGAGTGAATGTTCTATTGCTGAGAAACAGCTGAGAGTGTGGGTACTTTAGA 6436
Db 16390 GCCTGGGAGTGAATGTTCTATTGCTGAGAAACAGCTGAGAGTGTGGGTACTTTAGA 16449
QY 6437 CCTTTAGAGGCTGGCTTCACTAGCCCTGTGAGGTTTCTCTCTGAGTGGCAATGAGAT 6496
Db 16450 CCTTTAGAGGCTGGCTTCACTAGCCCTGTGAGGTTTCTCTCTGAGTGGCAATGAGAT 16509
QY 6497 ACCCTCCCTTGAACCCGTGGCATCAATTTGGTAAACCATCTTAATAATACCTAGGCTGT 6556
Db 16510 ACCCTCCCTTGAACCCGTGGCATCAATTTGGTAAACCATCTTAATAATACCTAGGCTGT 16569
QY 6557 TCTGAGTTGAGTGAAGCAATTAATAGTCTGACAGTGTGAGAAATATCCCAAGAGA 6616
Db 16570 TCTGAGTTGAGTGAAGCAATTAATAGTCTGACAGTGTGAGAAATATCCCAAGAGA 16629
QY 6617 GTGAGCAACCAATCACTCAACCTGAGATATATATTAATGAGCAAGTGTAAAGAT 6676
Db 16630 GTGAGCAACCAATCACTCAACCTGAGATATATATTAATGAGCAAGTGTAAAGAT 16689
QY 6677 ATTAATTCGTGAAATATTTTTCACACAAATTTTGGCTCTGACCTTGGACAA 6736
Db 16690 ATTAATTCGTGAAATATTTTTCACACAAATTTTGGCTCTGACCTTGGACAA 16749
QY 6737 ATTTGACAGTATGACTATCAAGTTCTGTGAAATAATCATCACCAATGAGAGCAAA 6796
Db 16750 ATTTGACAGTATGACTATCAAGTTCTGTGAAATAATCATCACCAATGAGAGCAAA 16809
QY 6797 TCTCCACAGCAGGATTTGACACTATATTAAGAAACATACAGCTAAGTGAACACACCT 6856
Db 16810 TCTCCACAGCAGGATTTGACACTATATTAAGAAACATACAGCTAAGTGAACACACCT 16869
QY 6857 GTTGTGAAATATCAATTTAACTGAGAACTTACGCTATGTGAACACATTAAGTATCA 6916
Db 16870 GTTGTGAAATATCAATTTAACTGAGAACTTACGCTATGTGAACACATTAAGTATCA 16929
QY 6917 AGAGAAACACAGCCATGTGGGAGGCCATTTGGGAGAGACACAGACAAAGTGAATGCA 6976
Db 16930 AGAGAAACACAGCCATGTGGGAGGCCATTTGGGAGAGACACAGACAAAGTGAATGCA 16989
QY 6977 GAAAGAGAGAGAGTGAAGTGAAGATTTGAAAACAGGGCCACAGAGAAACACAGAAA 7036
Db 16990 GAAAGAGAGAGAGTGAAGTGAAGATTTGAAAACAGGGCCACAGAGAAACACAGAAA 17049
QY 7037 TAGAGAGACACCAAGCCATCTAGAGATCAAGAACTTCAATGGCCATGTGGCATTAATG 7096
Db 17050 TAGAGAGACACCAAGCCATCTAGAGATCAAGAACTTCAATGGCCATGTGGCATTAATG 17109
QY 7097 AGAATGCTACTGAACCTTAATGAAAATGTCAATGATGTTCCATAGCTGTGAGAGAG 7156
Db 17110 AGAATGCTACTGAACCTTAATGAAAATGTCAATGATGTTCCATAGCTGTGAGAGAG 17169
QY 7157 CCCACAGCATGAGAGAAACCTTATATTAATAATCCAGGCGGGCGTGTAGTCAAC 7216
Db 17170 -CCACAGCATGAGAGAAACCTTATATTAATAATCCAGGCGGGCGTGTAGTCAAC 17228
QY 7217 GCGTGTAAATCTTAGACCTTTGGAGGCTGAGGAGGTGATGTCTTGAAGCGTTGAGCC 7276
Db 17229 GCGTGTAAATCTTAGACCTTTGGAGGCTGAGGAGGTGATGTCTTGAAGCC -GCTTGAAGCC 17287
QY 7277 TAGAGTTTGAAGCAGCGCTGGGCAACATGGCAAAACCTCATCTCAAAAAATTAATA 7336
Db 17288 TAGAGTTTGAAGCAGCGCTGGGCAACATGGCAAAACCTCATCTCAAAAAATTAATA 17347
QY 7337 AATTAGTCGGGTGTGTGATGCGTTCTTATAGTCCCATTACTTCAAGAGCTGAGCCGG 7396
Db 17348 AATTAGTCGGGTGTGTGATGCGTTCTTATAGTCCCATTACTTCAAGAGCTGAGCCGG 17406
QY 7397 AAGGTGAGGCTTCAATGAGCCGTGATGTGTCTACTGACCTCAAGCTGGGTGACAGAGT 7456
Db 17407 AAGGTGAGGCTTCAATGAGCCGTGAT -GATCTACTGACCTCAAGCTGGGTGACAGAGT 17465

OY	3077	TTCCATGACCCTGATCAAGGCTCTCATAGGGCAAGGAAGTGTCCAAACAATG	3136
Db	13091	TTCCATGACCCTGATCAAGGCTCTCATAGGGCAAGGAAGTGTGTCCAAACAATG	13150
OY	3137	GCACCTGGCTGGAAGAGCATGTGGAGCATGAATCCAGCTGTGCTGCTCAACAATGCC	3196
Db	13151	GCACCTGGCTGGAAGAGCATGTGGAGCATGAATCCAGCTGTGCTGTCCAAATGCC	13210
OY	3197	ATGTGACCCAGGCTGTGTCTGTCTCAGAGGAAGGGGTGCCCTTTTCTTACCTTGCCAAGGT	3256
Db	13211	ATGTGACCCAGGCTGTGTCTGTCTCAGAGGAAGGGGTGCCCTTTTCTTACCTTGCCAAGGT	13270
OY	3257	GCTGTGTGTGGGGGAAGTCCCTGACCTGTGGCTTTGTTTTCCCTCGCTCTTTTCTCT	3316
Db	13271	GCTGTGTGTGGGGGAAGTCCCTGACCTGTGGCTTTGTTTTCCCTCGCTCTTTTCTCT	13330
OY	3317	CTCTTCTCAAAATGTCTCAATCTATCTCAACAAGTCCCTATGTCTCTTGGGATCATC	3376
Db	13331	CTCTTCTCAAAATGTCTCAATCTATCTCAACAAGTCCCTATGTCTCTTGGGATCATC	13390
OY	3377	CTAGCCTTTCATATACCTTCCCTCAGTGATCTCAACCATACCTTGGCTTGAGGAATA	3436
Db	13391	CTAGCCTTTCATATACCTTCCCTCAGTGATCTCAACCATACCTTGGCTTGAGGAATA	13450
OY	3437	TCATNGCTGNGGGAACATGAGATCTATCTTCTTGAAGTCCAGACATCTCTTCCAT	3496
Db	13451	TCATNGCTGNGGGAACATGAGATCTATCTTCTTGAAGTCCAGACATCTCTTCCAT	13510
OY	3497	TGTATGTTCTACAGGACCTTAAATTCAGATCCCCCAACTAAGCTTTCATCTTCTT	3556
Db	13511	TGTATGTTCTACAGGACCTTAAATTCAGATCCCCCAACTAAGCTTTCATCTTCTT	13570
OY	3557	ACAAACCAACCTTTCCTGTGTGTTTCTGTATTCAGTAAATGACCCCAAAATGTGCTGA	3616
Db	13571	ACAAACCAACCTTTCCTGTGTGTTTCTGTATTCAGTAAATGACCCCAAAATGTGCTGA	13630
OY	3617	TTACTTACAAACCAAGTGCACACAGGGTCTATGATCTGGGCTTGTGTATCTTCTAGGT	3676
Db	13631	TTACTTACAAACCAAGTGCACACAGGGTCTATGATCTGGGCTTGTGTATCTTCTAGGT	13690
OY	3677	TTATCTCCTCCCTGCACAACTTCACTGTGTCCAGCATACGAAATCTACATGAGTTTGA	3736
Db	13691	TTATCTCCTCCCTGCACAACTTCACTGTGTGTCCAGCATACGAAATCTACATGAGTTTGA	13750
OY	3737	GCACATGCTTCTCATGTGTTGGGCTTGCATGCTGCTCCCTGTGCTGAACAACCTTT	3796
Db	13751	GCACATGCTTCTCATGTGTTGGGCTTGCATGCTGCTCCCTGTGCTGAACAACCTTT	13810
OY	3797	CCTCACTTGTCAACCTGGAAATTCCTGTGATTTTTCAGCTCTTGGGCCAATGCTTCC	3856
Db	13811	CCTCACTTGTCAACCTGGAAATTCCTGTGATTTTTCAGCTCTTGGGCCAATGCTTCC	13870
OY	3857	TCTTGTGTGGAACCTTCCACAACCTCTCTAGGAGACCTTAGGACCTCTGTCTAATTC	3916
Db	13871	TCTTGTGTGGAACCTTCCACAACCTCTCTAGGAGACCTTAGGACCTCTGTCTAATTC	13930
OY	3917	TCACTGCACTTCTTACACTACACTTGGTGTGCAATGCTAGATTTGACAGAGTCTTT	3976
Db	13931	TCACTGCACTTCTTACACTACACTTGGTGTGCAATGCTAGATTTGACAGAGTCTTT	13990
OY	3977	CTGCTTTTGTACAGTGAACCTTCTGAAATGAAAGACAGATCTTGTATCTTCAAGTCT	4036
Db	14051	CTGCTTTTGTACAGTGAACCTTCTGAAATGAAAGACAGATCTTGTATCTTCAAGTCT	14110
OY	4037	CTGCAATGCTGGCAATATGATGTTATTAAGAGACGTGTTCTTGTGATGAAGATGAT	4096
Db	14051	CTGCAATGCTGGCAATATGATGTTATTAAGAGACGTGTTCTTGTGATGAAGATGAT	14110
OY	4097	GAATTAATTAATGAAGAAATGAATGAAGAAATTAACATATGGGTGATTTGACAGATGAACAG	4156
Db	14111	GAATTAATTAATGAAGAAATGAATGAAGAAATTAACATATGGGTGATTTGACAGATGAACAG	14170
OY	4157	TTGTGATATGTTTGTCAACATGATAGTGTTCAGATTAATGTGCCACAGAGTGTCTG	4216

Db	14171	TTGGATATATGTTTGTCAACCTGATGATGTTGCAGATTAATGTCACAGAGTGTCTG	14230
OY	42127	GGTACAGAGCTAGAGGCAATGTGTTATATGTAATAGTGACTGGATTTGGCAAACTGAGA	4276
Db	14231	GGTACAGAGCTAGAGGCAATGTGTTATATGTAATAGTGACTGGATTTGGCAAACTGAGA	14290
OY	4277	GTGTGTATATGTCAAAAGACAGCACTTGTGTCCAGATGAGCTGGAATGTGTAGG	4336
Db	14291	GTGTGTATATGTCAAAAGACAGCACTTGTGTCCAGATGAGCTGGAATGTGTAGG	14350
OY	4337	GCCACAGAGGATATGSTATPAACAGATGATPAATAAATGTGTAAATGACAGTGGC	4396
Db	14351	GCCACAGAGGATATGSTATPAACAGATGATPAATAAATGTGTAAATGACAGTGGC	14410
OY	4397	AGTATCTGGGGATGCAAGTCAGTCAAAAAGAGTACTTTTGAATGACAGGGGACAAAGTCTG	4456
Db	14411	AGTATCTGGGGATGCAAGTCAGTCAAAAAGAGTACTTTTGAATGACAGGGGACAAAGTCTG	14470
OY	4457	GGTATACCTCTCTGAAAAGAGAGAAAGATACCCAAAGTTGCTCAAGATGAATTTCC	4516
Db	14471	GGTATACCTCTCTGAAAAGAGAGAAAGATACCCAAAGTTGCTCAAGATGAATTTCC	14530
OY	4517	TGGAATCCCAATCCCACTGATGACACATGTGAAATGTCAACGTGAGAAATGTCAAGCTGA	4576
Db	14531	TGGAATCCCAATCCCACTGATGACACATGTGAAATGTCAACGTGAGAAATGTCAAGCTGA	14590
OY	4577	AGGCGTCTCTCTCTTCCCAAGAGTCAGTCCAGGCTGAAATCAAGAACTGAGATATGTGTG	4636
Db	14591	AGGCGTCTCTCTCTTCCCAAGAGTCAGTCCAGGCTGAAATCAAGAACTGAGATATGTGTG	14650
OY	4637	GGTTGTAGTATGACCAAGTATGATGTGTATATGTGTATGATGTGTGTGTGTGAGCCAC	4696
Db	14651	GGTTGTAGTATGACCAAGTATGATGTGTATATGTGTATGATGTGTGTGTGTGAGCCAC	14710
OY	4697	TATGTGCTTTTATATAGGAAGGAAATATAGGAAGCAACGATATGTCTATAGAGTCAAGTAG	4756
Db	14711	TATGTGCTTTTATATAGGAAGGAAATATAGGAAGCAACGATATGTCTATAGAGTCAAGTAG	14770
OY	4757	GGAAGAGAGGAAGTCAATCTGTGTGAAGCAACTGTGTATTAATCTGTGTAAATGTGCC	4816
Db	14771	GGAAGAGAGGAAGTCAATCTGTGTGAAGCAACTGTGTATTAATCTGTGTAAATGTGCC	14830
OY	4817	TCATTTGTTCCCATGACCCCTTAGAGCAAACTACCCCTTTAGTCTTACATCAAGTCTC	4876
Db	14831	TCATTTGTTCCCATGACCCCTTAGAGCAAACTACCCCTTTAGTCTTACATCAAGTCTC	14890
OY	4877	TCCATGGCCAAATCCCTATTTGGCCCTTCAGCTTTGACTTTATTAATCTTTAATCCTTAAC	4936
Db	14891	TCCATGGCCAAATCCCTATTTGGCCCTTCAGCTTTGACTTTATTAATCTTTAATCCTTAAC	14950
OY	4937	ACTAAGCTCCAGAAACCCATATGCTATTCCTGTACACTCAGTTTCTCATGCTCTGTGAA	4996
Db	14951	ACTAAGCTCCAGAAACCCATATGCTATTCCTGTACACTCAGTTTCTCATGCTCTGTGAA	15010
OY	4997	TCTTTCTCTCTCTGGGGTTCATCTCTCTCTGTGTGCTTTTAATCTCTACCTTCAGATT	5056
Db	15011	TCTTTCTCTCTCTGGGGTTCATCTCTCTCTGTGTGCTTTTAATCTCTACCTTCAGATT	15070
OY	5057	TCACCTTAAGTATCATCTTCCCTGGGAGTTTCCAGACTCTCCCACTGCTCTTGTCTG	5116
Db	15071	TCACCTTAAGTATCATCTTCCCTGGGAGTTTCCAGACTCTCCCACTGCTCTTGTCTG	15130
OY	5117	AGCTGATCCGATGTGTTTGGCTGCAATTTTGGTATGATCAACCCCTTTAGCCATC	5176
Db	15131	AGCTGATCCGATGTGTTTGGCTGCAATTTTGGTATGATCAACCCCTTTAGCCATC	15190
OY	5177	TCTCTGATGCTGTGAGCTCAATGTGTCAAGTACATTAATCTGAGCCATCTCTGGACCCA	5236
Db	15191	TCTCTGATGCTGTGAGCTCAATGTGTCAAGTACATTAATCTGAGCCATCTCTGGAGTCCA	15250
OY	5237	GAGAAAGCAAAAGAGGGCTGATACCCGGCTTCACCAAATGCTGTGATTGATTGACA	5296

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 148198)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On Jun 14, 2000 this sequence version replaced gi:7715652.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0223K12
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 142463 bases at least Q40
Consensus quality: 144172 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 15100; agarose-fp
Insert size: 147398; sum-of-contigs
Quality coverage: 4.53 in Q20 bases; sum-of-contigs
Quality coverage: 4.67 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2755: contig of 2755 bp in length
2756 2855: gap of unknown length
2856 12253: contig of 9398 bp in length
12254 12353: gap of unknown length
12354 22567: contig of 10214 bp in length
22568 22667: gap of unknown length
22668 37602: contig of 14935 bp in length
37603 37702: gap of unknown length
37703 51053: contig of 13351 bp in length
51054 51153: gap of unknown length
51154 66693: contig of 15540 bp in length
66694 66793: gap of unknown length
66794 85880: contig of 19087 bp in length
85881 85980: gap of unknown length
85981 115955: contig of 29975 bp in length
115956 148198: contig of 32143 bp in length.

FEATURES
source
1. 148198
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-223K12"
1. 2755
/note="assembly_name:Contig4"
2856. 12253
/note="assembly_name:Contig5"
12354. 22567
/note="assembly_name:Contig6"
22668. 37602
/note="assembly_name:Contig7"
37703. 51053
/note="assembly_name:Contig8"
51154. 66693
/note="assembly_name:Contig9"

misc_feature
66794..85880
/note="assembly_name:Contig10"
85981..115955
/note="assembly_name:Contig11"
clone_end:SP6
vector_side:right"
misc_feature
116056..148198
/note="assembly_name:Contig12"
BASE COUNT 38577 a 34635 c 33776 g 40404 t 806 others
ORIGIN
Query Match 75.6%; Score 7556.4; DB 2; Length 148198;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 7636; Conservative 0; Mismatches 121; Indels 7; Gaps 6;
QY 2237 TAGTGCATGGAATGTTGCTGAAGCGGACCAAGCCAGACCGTGA 2296
12251 TGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 12310
QY 2297 GGGCCGACTCCCTAAGCTGCTTATCTGCTGTTGCTCTGCTTATCACCCT 2356
12311 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 12370
QY 2357 CTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTGTGCACTCCCTACCCCAAGAC 2416
12371 CTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTGTGCACTCCCTACCCCAAGAC 12430
QY 2417 GCTGCTCAGATGCTGAGAGCTTTGATGCTACCACTGATGACAAATGGAACATGCTGT 2476
12431 GCTGCTCAGATGCTGAGAGCTTTGATGCTACCACTGATGACAAATGGAACATGCTGT 12490
QY 2477 TTTTAAAGGTAGAGAGGAGTGAAGTTAGGCGTTTGAAGCTTGAAGCTTCTCTTC 2536
12491 TTTTAAAGGTAGAGAGGAGTGAAGTTAGGCGTTTGAAGCTTGAAGCTTCTCTTC 12550
QY 2537 ACAAGGGTGTCCCTGCTGTGAGAGGCTTAGAATTTATCTATGATGATCACTGACAGC 2596
12551 ACAAGGGTGTCCCTGCTGTGAGAGGCTTAGAATTTATCTATGATGATCACTGACAGC 12610
QY 2597 TTCTCTCAAGCTATCTCAGTAGTCAAGGTTTCTCACTGAGCCCTCAGTAGTGTGG 2656
12611 TTCTCTCAAGCTATCTCAGTAGTCAAGGTTTCTCACTGAGCCCTCAGTAGTGTGG 12670
QY 2657 TTTTTCAGGGAGTTTGTGTGGAAGTCAAAATGGGACCGGAGTTATCTCAGAGA 2716
12671 TTTTTCAGGGAGTTTGTGTGGAAGTCAAAATGGGACCGGAGTTATCTCAGAGA 12730
QY 2717 GATGGAAGATTTCCCGAGCCCTGTGAGTCTGATTCGCTCAAGGTCAACAGTGTCT 2776
12731 GATGGAAGATTTCCCGAGCCCTGTGAGTCTGATTCGCTCAAGGTCAACAGTGTCT 12790
QY 2777 TTCTGATCAAGTACTGCTGGGCCCAAAATCAGGCGCCAGGCTGGAAGGCTGGAATGAC 2836
12791 TTCTGATCAAGTACTGCTGGGCCCAAAATCAGGCGCCAGGCTGGAAGGCTGGAATGAC 12850
QY 2837 ACTGGGAGCCCTTCCCAATAGGCTTGGCATGAGGCCATATAGATAGTATGATGATG 2896
12851 ACTGGGAGCCCTTCCCAATAGGCTTGGCATGAGGCCATATAGATAGTATGATGATG 12910
QY 2897 TCTTTCCAGTGTCCCTCTTCTCTGTAAGAGCTGGGCTTAAGGAGTGTGATGCGTGT 2956
12911 TCTTTCCAGTGTCCCTCTTCTCTGTAAGAGCTGGGCTTAAGGAGTGTGATGCGTGT 12970
QY 2957 GGGCTGGAAGTGAACATCAATGAGTCTTCTTCACTAGTCTTATGATGATGATGATG 3016
12971 GGGCTGGAAGTGAACATCAATGAGTCTTCTTCACTAGTCTTATGATGATGATGATG 13030
QY 3017 CTCTGCTGTAACCTCAACATGCTGGCCATCTTTTCTTCTGCAATGCTTCTCTTGGCC 3076
13031 CTCTGCTGTAACCTCAACATGCTGGCCATCTTTTCTTCTGCAATGCTTCTCTTGGCC 13090

Db 23737 CTTTGGCTCTTCAAACTTGGCTGCACTGAGCAATCATCTGCTCTCCACTTTCAC 2378
 QY 8089 TACTCATGATCTTTGAGTCTGAGTCCAACTGATACCTCCCTCAGTGTCTTTTTC 8148
 Db 23677 TACTCATGATCTTTGAGTCTGAGTCCAACTGATACCTCCCTCAGTGTCTTTTTC 23618
 QY 8149 TAGTAAGATTTCCAGAGAGGGAATCTGAATGAGCCAGTCCATATTTTTCAGACCAAC 8208
 Db 23617 TAGTAAGATTTCCAGAGAGGGAATCTGAATGAGCCAGTCCATATTTTTCAGACCAAC 23558
 QY 8209 ATTAAAGGTGTTATTTGAGTCTGAGTCCAACTGATACCTCCCTCAGTGTCTTTTTC 8268
 Db 23557 ATTAAAGGTGTTATTTGAGTCTGAGTCCAACTGATACCTCCCTCAGTGTCTTTTTC 23498
 QY 8269 TTTACTTCAATGCAAAAGCAAGATAGCTGTGTTCTCAAAATAGGGCCCTGAGGCAAG 8328
 Db 23497 TTTACTTCAATGCAAAAGCAAGATAGCTGTGTTCTCAAAATAGGGCCCTGAGGCAAG 23438
 QY 8329 TGTGTGCTCATGCTATATATCCCAACTGTGAGAGCCGAGAGGAGGAGATCACTTG 8388
 Db 23437 TGTGTGCTCATGCTATATATCCCAACTGTGAGAGCCGAGAGGAGGAGATCACTTG 23378
 QY 8389 AGTCCAGAGTCTTGAACCAAGCTGAGGCAATGTGGAATCTCATCTTAATAAATA 8448
 Db 23377 AGTCCAGAGTCTTGAACCAAGCTGAGGCAATGTGGAATCTCATCTTAATAAATA 23318
 QY 8449 CAAAAAATTAAGCAGAGTGTGAGTGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 8508
 Db 23317 CAAAAAATTAAGCAGAGTGTGAGTGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 23258
 QY 8509 GTGGAGAGATTTGCTCAGTGTGAGGAGGAGATTTGCAAGTGAACCTGAGTGTGCTG 8568
 Db 23257 GTGGAGAGATTTGCTCAGTGTGAGGAGGAGATTTGCAAGTGAACCTGAGTGTGCTG 23198
 QY 8569 CAATCAAGCTGAGTGTGCAAGATTTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 8628
 Db 23197 CAATCAAGCTGAGTGTGCAAGATTTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 23138
 QY 8629 TAAATATGTTCTGAGCAGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 8688
 Db 23137 TAAATATGTTCTGAGCAGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 23078
 QY 8689 TCAGTCAAGCTGTATATCACTGATCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 8748
 Db 23077 TCAGTCAAGCTGTATATCACTGATCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 23018
 QY 8749 CACTAGATTTTATCACTGATCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 8808
 Db 23017 CACTAGATTTTATCACTGATCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22958
 QY 8809 AGGCACTGAGATTAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 8868
 Db 22957 AGGCACTGAGATTAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22898
 QY 8869 AAAGTGAAGAGCAATTAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 8928
 Db 22897 AAAGTGAAGAGCAATTAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22838
 QY 8929 AAAGTGTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 8988
 Db 22837 AAAGTGTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22778
 QY 8989 CTTTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9048
 Db 22777 CTTTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22718
 QY 9049 TGTGTGAGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9108
 Db 22717 TGTGTGAGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22658
 QY 9109 GGAAGCTGTGTGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9168
 Db 22657 GGAAGCTGTGTGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22598

QY 9169 AGGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9228
 Db 22597 AGGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22538
 QY 9229 AGAAGTCTTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9288
 Db 22537 AGAAGTCTTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22478
 QY 9289 ACTTTTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9348
 Db 22477 ACTTTTGAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22418
 QY 9349 ATCCGTGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9408
 Db 22417 ATCCGTGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22358
 QY 9409 GGTGTGAGGAATCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9468
 Db 22357 GGTGTGAGGAATCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22298
 QY 9469 TAAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9528
 Db 22297 TAAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22238
 QY 9529 GAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9588
 Db 22237 GAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22178
 QY 9589 CTCTACTAATAAATACAAATTAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9648
 Db 22177 CTCTACTAATAAATACAAATTAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22118
 QY 9649 TTGGAGGCTGTGAGGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9708
 Db 22117 TTGGAGGCTGTGAGGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 22058
 QY 9709 GATTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9768
 Db 22057 GATTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 21998
 QY 9769 AATGATGATGATATACAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9828
 Db 21997 AATGATGATGATATACAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 21938
 QY 9829 CTTTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9888
 Db 21937 CTTTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 21878
 QY 9889 AGGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 9948
 Db 21877 AGGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 21818
 QY 9949 CTATTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 10000
 Db 21817 CTATTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 21766

RESULT 3
 AC024232 148198 bp DNA linear HNG 07-JUL-2000
 LOCUS Homo sapiens chromosome 11, clone RP11-223K12, WORKING DRAFT
 DEFINITION
 SEQUENCE, 9 unordered pieces.
 ACCESSION AC024232
 VERSION AC024232.3 GI:8516169
 KEYWORDS HNG; HNGS PHASE1; HNGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 148198)
 AUTHORS Waterston R.H.
 TITLE The sequence of Homo sapiens clone

Db 28117 GATCGGCGCTTGTATCTTCTCAGGTTATCTCTCCCTCCGACATTCAGTGTGTC 28058
Qy 3709 CAGCCATAGAAATCTATCAGATGGTTGGAGCACTGCTTCTCTCAATGTTGGGCTCGCAT 3768
Db 28057 CAGCCATAGAAATCTATCAGATGGTTGGAGCACTGCTTCTCTCAATGTTGGGCTCGCAT 27998
Qy 3769 GCTGCTCCCTCTCTCTGTGTAAACACCTTCTCTCACTTGTCAACCTGTAAATTCCTGTGA 3828
Db 27997 GCTGCTCCCTCTCTGTGTAAACACCTTCTCTCACTTGTCAACCTGTAAATTCCTGTGA 27938
Qy 3829 TTTTTCAGCTCTTGGGCGCCCAATGCTTCTCTTGTGTGAACCTTCCACAATCTTCTGA 3888
Db 27937 TTTTTCAGCTCTTGGGCGCCCAATGCTTCTCTTGTGTGAACCTTCCACAATCTTCTGA 27878
Qy 3889 GGGAGACTTAGGACCTGTCTATATCTCAGTGCCTCTTCACTCAACCTTGTGAGT 3948
Db 27877 GGGAGACTTAGGACCTGTCTATATCTCAGTGCCTCTTCACTCAACCTTGTGAGT 27818
Qy 3949 TGCATGGCTAGATTCAGAGAGTCTTCTGCTTTTGTACAGTGAATCTTCTGAAGTGA 4008
Db 27817 TGCATGGCTAGATTCAGAGAGTCTTCTGCTTTTGTACAGTGAATCTTCTGAAGTGA 27758
Qy 4009 AGACAGAGTCTTGTATCTCTCAGTGCCTCTCAACATGCTGGACATATAGTATTCAG 4068
Db 27757 AGACAGAGTCTTGTATCTCTCAGTGCCTCTCAACATGCTGGACATATAGTATTCAG 27698
Qy 4069 TGACTGTTCTTGTAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 4128
Db 27697 TGACTGTTCTTGTAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 27638
Qy 4129 AACGTATGGTGTATTCAGAGATGAACAGTGTGAGATATGTTGTCAACATGATAGTGT 4188
Db 27637 AACGTATGGTGTATTCAGAGATGAACAGTGTGAGATATGTTGTCAACATGATAGTGT 27578
Qy 4189 GCGATTAATATGTCACAGAGATGTCGTGGTACAGAGTACAGAGCATGTGTATATGA 4248
Db 27577 GCGATTAATATGTCACAGAGATGTCGTGGTACAGAGTACAGAGCATGTGTATATGA 27518
Qy 4249 ATAGTACTGGATTTGTGACAACTGAGAGTGTATATGTCAAAGAGACAGCATTTGT 4308
Db 27517 ATAGTACTGGATTTGTGACAACTGAGAGTGTATATGTCAAAGAGACAGCATTTGT 27458
Qy 4309 GTCACAGATGAGTGAATGTGTAGGGGACAGAGAGATATGTTATAGCACATGTA 4368
Db 27457 GTCACAGATGAGTGAATGTGTAGGGGACAGAGAGATATGTTATAGCACATGTA 27398
Qy 4369 TAAAAAATGTGTAAATGACAGATGACATATCTGGGAGATGACATCAAAAGAGAT 4428
Db 27397 TAAAAAATGTGTAAATGACAGATGACATATCTGGGAGATGACATCAAAAGAGAT 27338
Qy 4429 ACTTTTGAATGACAGGGGACAAAGTCTGGGTATACCTCTCTGAAAAGAGAGAAAGAT 4488
Db 27337 ACTTTTGAATGACAGGGGACAAAGTCTGGGTATACCTCTCTGAAAAGAGAGAAAGAT 27278
Qy 4489 ACCCAAGTTGCTCAAGATGAATTTCTGTGAATCCCATCCCACTGGATGACGCTGTG 4548
Db 27277 ACCCAAGTTGCTCAAGATGAATTTCTGTGAATCCCATCCCACTGGATGACGCTGTG 27218
Qy 4549 AATGTACCGTGAAGATGTCAAGCTGAAGCGTCTCTTCTTCAAGATCAGTCCAGGC 4608
Db 27217 AATGTACCGTGAAGATGTCAAGCTGAAGCGTCTCTTCTTCAAGATCAGTCCAGGC 27158
Qy 4609 TGGAAATCCAAAGACCTGAGATAGTGTGGGTGTGTGTGTGTGTGTGTGTGTGTGT 4668
Db 27157 TGGAAATCCAAAGACCTGAGATAGTGTGGGTGTGTGTGTGTGTGTGTGTGTGTGT 27098
Qy 4669 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4728
Db 27097 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 27038
Qy 4729 GCCACATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4788
Db 27037 GCCACATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 26978

Qy 4789 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4848
Db 26977 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 26918
Qy 4849 ACCCTCTTATAGTCTTATCAATCAAGTCTTCTCATAGGCAATTCCTTATTTGGCCCTTCAGCT 4908
Db 26917 ACCCTCTTATAGTCTTATCAATCAAGTCTTCTCATAGGCAATTCCTTATTTGGCCCTTCAGCT 26858
Qy 4909 TTTGACTTATATATATCTTTTAACTTAACTAAGTCTCCAGAAACCTTATGCTATTTCTGTG 4968
Db 26857 TTTGACTTATATATATCTTTTAACTTAACTAAGTCTCCAGAAACCTTATGCTATTTCTGTG 26798
Qy 4969 TACACTCAGTTTGTCTCAAGTCTTGTGAATCTTCTCTCTGTGGGTTCTCATCTCTCT 5028
Db 26797 TACACTCAGTTTGTCTCAAGTCTTGTGAATCTTCTCTCTCTGTGGGTTCTCATCTCTCT 26738
Qy 5029 GTGTGCTTTTAAATTCCTTATCTTCAATTCAGATTCATCTTAAATCATCTTCCCTGGAAATTT 5088
Db 26737 GTGTGCTTTTAAATTCCTTATCTTCAATTCAGATTCATCTTAAATCATCTTCCCTGGAAATTT 26678
Qy 5089 TCCCAAGCTCTCCCACTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5148
Db 26677 TCCCAAGCTCTCCCACTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 26618
Qy 5149 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5208
Db 26617 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 26558
Qy 5209 ACCATATATCTGGCCATCTCTGGAGACCAAGAAAGCAAAAGAGGCGGTAAACCGGTCT 5268
Db 26557 ACCATATATCTGGCCATCTCTGGAGACCAAGAAAGCAAAAGAGGCGGTAAACCGGTCT 26498
Qy 5269 CACCAATATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5328
Db 26497 CACCAATATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 26438
Qy 5329 ACCGGAACATGAAGAGAGGTTCTGTGGCAGCTGTGTGGAACTGTCTGTGCTGTGAGA 5388
Db 26437 ACCGGAACATGAAGAGAGGTTCTGTGGCAGCTGTGTGGAACTGTCTGTGCTGTGAGA 26378
Qy 5389 TGGCTGGGCGGCTACTATCTGCTTCAAGGTTAAACCAATTCCTGTGCTTCAACCTGTCAAG 5448
Db 26377 TGGCTGGGCGGCTACTATCTGCTTCAAGGTTAAACCAATTCCTGTGCTTCAACCTGTCAAG 26318
Qy 5449 GGAAGATGCTTCCAGATACCGCGGGAATGTCGAGACTTCAATGCTGCTGCTGCTGCTGCT 5508
Db 26317 GGAAGATGCTTCCAGATACCGCGGGAATGTCGAGACTTCAATGCTGCTGCTGCTGCTGCT 26258
Qy 5509 AGAGGTGAAGAAAGCCCTAGACATGAGACCTGTCAAGATTCATGCACTTCCCTGAGCTT 5568
Db 26257 AGAGGTGAAGAAAGCCCTAGACATGAGACCTGTCAAGATTCATGCACTTCCCTGAGCTT 26198
Qy 5569 GTGATCTCAGTGTCTGAGTCTCACTTAACTCCGTGTGGACAACCTTGGCCCTTAA 5628
Db 26197 GTGATCTCAGTGTCTGAGTCTCACTTAACTCCGTGTGGACAACCTTGGCCCTTAA 26138
Qy 5629 TCTAGCCCAATTCATCTGTGATTTTCCATTTGCCATTTGATGGGAAACCCACACCCC 5688
Db 26137 TCTAGCCCAATTCATCTGTGATTTTCCATTTGCCATTTGATGGGAAACCCACACCCC 26078
Qy 5689 ACTAAAGCCAGCATCTTCAACCTTGAACCTGACATCTGACCTCTGGCTCTCTGTGTG 5748
Db 26077 ACTAAAGCCAGCATCTTCAACCTTGAACCTGACATCTGACCTCTGGCTCTCTGTGTG 26018
Qy 5749 TTTCTCTCAACCAATTTCTCTTCCAGGCAATGGAACAAGAAATGGCATGGCAATGGCAAT 5808
Db 26017 TTTCTCTCAACCAATTTCTCTTCCAGGCAATGGAACAAGAAATGGCATGGCAATGGCAAT 25958
Qy 5809 CAGTACCAACCAATGGCCGTGAGTATATGCGCTGTAGCCCAATCTAGTCTTGTCTGCACT 5868
Db 25957 CAGTACCAACCAATGGCCGTGAGTATATGCGCTGTAGCCCAATCTAGTCTTGTCTGCACT 25898

QY	1501	TTCTCTAGAGTATTTCTTAAGTAAGTTTATCTGTGTGTGGTTACTTAAAGTAAGTATAT	1560
Db	30277	TTCTCTAGTATTTTCTTAAGTAAGTTTATCTGTGTGTGGTTACTTAAAGTAAGTATAT	30211
QY	1561	TGCTGTCTTAGTGTGCTTCTTAGAGAGAGGCTTGAACAAGTACAAGCAATATTAGTTCCT	1620
Db	30217	TGCTGTCTCTTAGTGTGCTTCTTAGAGAGAGGCTTGAACAAGTACAAGCAATATTAGTTCCT	30155
QY	1621	CTTTTCTCACTCTCCCCATTTGTGAGATTAACCTCAATCACAAAAGGTGATCCTCACTCT	1680
Db	30157	CTTTTCTCACTCTCCCCATTTGTGAGATTAACCTCAATCACAAAAGGTGATCCTCACTCT	30099
QY	1681	ACTCACTTCCCTGACTTATGGAATCCCTGGAACCCATTCGCACTGTGAAGTCAACAGCTGGA	1740
Db	30097	ACTCACTTCCCTGACTTATGGAATCCCTGGAACCCATTCGCACTGTGAAGTCAACAGCTGGA	30033
QY	1741	CGTCAGCAGTGTACCCAGCTTACTGCTTGAATAATTCGTGAAGGGGTTGGGGGCACTGTG	1800
Db	30037	CGTCAGCAGTGTACCCAGCTTACTGCTTGAATAATTCGTGAAGGGGTTGGGGGCACTGTG	29977
QY	1801	CCGGGAAAAGAGACTTTTGATTCAGATTTCTGTCCAGACCCTGACCTTATTGCACTGA	1860
Db	29977	CCGGGAAAAGAGACTTTTGATTCAGATTTCTGTCCAGACCCTGACCTTATTGCACTGA	29911
QY	1861	TGTATATCAGCCCAATATTGGCTTACTCCTCGGAGACACACACTTCCCACTAGATTTGAGG	1920
Db	29917	TGTATATCAGCCCAATATTGGCTTACTCCTCGGAGACACACACTTCCCACTAGATTTGAGG	29855
QY	1921	TGGGGGTGTGTCTGTCTGTCAACTCTATATAGGAGTTCAACTGTGTCAACCCAGACTGTCC	1980
Db	29857	TGGGGGTGTGTGTCTGTCTGTCAACTCTATATAGGAGTTCAACTGTGTCAACCCAGACTGTCC	29799
QY	1981	TGTGGCCTCTTGACAGCTCAGCATGGCTAGGGTACTGGAGACACCCTTGCACTGGGGTTGT	2040
Db	29797	TGTGGCCTCTTGACAGCTCAGCATGGCTAGGGTACTGGAGACACCCTTGCACTGGGGTTGT	29733
QY	2041	GGAGCCTTATGCTGTCTTGTGGCAATGCGCACCCCTCTTCTCGTAGATTAAGCTGGGAC	2100
Db	29737	GGAGCCTTATGCTGTCTTGTGGCAATGCGCACCCCTCTTCTCGTAGATTAAGCTGGGAC	29677
QY	2101	TAGAAGCGAAGATTAAGTTCTGGGCTAGGGTTAAGTGAAGGCGAGTTTAAAGGCTGGT	2160
Db	29677	TAGAAGCGAAGATTAAGTTCTGGGCTAGGGTTAAGTGAAGGCGAGTTTAAAGGCTGGT	29611
QY	2161	CAAAATTTGGGGTCAAGGGCTATGCGAAAGGAGTCGCTCCCAATGATCAAGATATCTATT	2220
Db	29617	CAAAATTTGGGGTCAAGGGCTATGCGAAAGGAGTCGCTCCCAATGATCAAGATATCTATT	29555
QY	2221	TTGTTCCTCCCTAGAGACTATGTGCCCATATGCGCAATGGGAATGTTGCTGAAGCGCAAGCCAGACC	2280
Db	29557	TTGTTCCTCCCTAGAGACTATGTGCCCATATGCGCAATGGGAATGTTGCTGAAGCGCAAGCCAGACC	29499
QY	2281	CAGACGTGACGTGTGAAGGCGCTGATCCCTTAAGTCTGTCTTATCTGTCTGTGTGTCTC	2340
Db	29497	CAGACGTGACGTGTGAAGGCGCTGATCCCTTAAGTCTGTCTTATCTGTCTGTGTGTCTC	29433
QY	2341	TGCAATTTATCACTCTGTG-----TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAACTT	2388
Db	29437	TGCAATTTATCACTCTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAACTT	29377
QY	2389	TGCGATCTCCCTAAGCTCAACCCAGAACGCTGTCTCAGATGCGCTGAGCTTATGATGTACC	2448
Db	29377	TGCGATCTCCCTAAGCTCAACCCAGAACGCTGTCTCAGATGCGCTGAGCTTATGATGTACC	29311
QY	2449	ACCGTGAATGACATNGAACATGCTGTTTTTAAAGTATAGAGAGGACTGAGGTTAAGGC	2508
Db	29317	ACCGTGAATGACATNGAACATGCTGTTTTTAAAGTATAGAGAGGACTGAGGTTAAGGC	29255
QY	2509	GTTTAGAGACTTAACTACTCTCTTCACAAAGGGTCTCCCTGTGTGGAGAGTCTTA	2568
Db	29257	GTTTAGAGACTTAACTACTCTCTTCACAAAGGGTCTCCCTGTGTGTGGAGAGTCTTA	29199
QY	2569	GGAATTAATCTATGCTATCACTGACAGCTTCTTCAAGCTATCTCAGTAGTCAAAAGTT	2628

Db	29197	GGAAATTACTGATGATATCATCTGACAGCTTCTCTCAAGCTATCTCAGTAGGTCAAAGTT	2913
QY	2629	TCTCACTGGGCCCCCTCAGTAGTGATGGGTTTTTTCAGGGAGTTTGTTGAAAGACTCAC	2688
Db	29137	TCTCACTGGGCCCCCTCAGTAGTGATGGGTTTTTTCAGGGAGTTTGTTGAAAGACTCAC	2907
QY	2689	AAATGGGACCGGAGTTTAATCTCAGAGAGATGGAGAAATTTCCCAAGCCTCTGGATGCT	2748
Db	29077	AAATGGGACCGGAGTTTAATCTCAGAGAGATGGAGAAATTTCCCAAGCCTCTGGATGCT	2901
QY	2749	GCATTCCGTAAAGGTCAACAAGTGCTTCTGATCAAGGATCTGCTGGGCCCCAAATCAG	2808
Db	29017	GCATTCCGTAAAGGTCAACAAGTGCTTCTGATCAAGGATCTGCTGGGCCCCAAATCAG	2895
QY	2809	GGCCAGGCTGGAAAGGCTGGAAATCGACACTGGGGAACCTTCCCAATATGACCTTTGGCA	2868
Db	28957	GGCCAGGCTGGAAAGGCTGGAAATCGACACTGGGGAACCTTCCCAATATGACCTTTGGCA	2889
QY	2869	TGAGCCCAATAGCAATATGATAGAGATTTCTTTCCATATGACCCTCTTCTCTGTAAAG	2928
Db	28897	TGAGCCCAATAGCAATATGATAGAGATTTCTTTCCATATGACCCTCTTCTCTGTAAAG	2883
QY	2929	CTTGGGCTTAAGGGAGTGTGATGTGTGTGGGCCCTGGAGGATGACCAATCAAGTGGCTTT	2988
Db	28837	CTTGGGCTTAAGGGAGTGTGATGTGTGTGGGCCCTGGAGGATGACCAATCAAGTGGCTTT	2877
QY	2989	CTTCAGTCTTAAGTCTTAAGTTCTACACCGGCTCTGCTGACTCAACACTGGGCATCTT	3048
Db	28777	CTTCAGTCTTAAGTCTTAAGTTCTACACCGGCTCTGCTGACTCAACACTGGGCATCTT	2871
QY	3049	TTTTTCTCTGGCAATGCTCTTCCCTTGCTTCATGACCTGTATCAAGTCTCTTCTTAAG	3108
Db	28717	TTTTTCTCTGGCAATGCTCTTCCCTTGCTTCATGACCTGTATCAAGTCTCTTCTTAAG	2865
QY	3109	GGCAAGGCAAGTGTGCCCAACAACAAGGACCGGCTTAAGAGCAATGTGGAGCATGAA	3168
Db	28657	GGCAAGGCAAGTGTGCCCAACAACAAGGACCGGCTTAAGAGCAATGTGGAGCATGAA	2859
QY	3169	ATCCAGTCTGTGTGCTCAACAAGTCCCATGTGACCCAGGCTGTGTGCTCAGAGAGAG	3228
Db	28577	ATCCAGTCTGTGTGCTCAACAAGTCCCATGTGACCCAGGCTGTGTGCTCAGAGAGAG	2853
QY	3229	GGGTGCTTTTCTTAACCTTGCCAAAGGTGTGTGTGGGAAAGTCTGACTGTGGC	3288
Db	28537	GGGTGCTTTTCTTAACCTTGCCAAAGGTGTGTGTGGGAAAGTCTGACTGTGGC	2847
QY	3289	TTTGTTTTCCCTCTGCGCTCTTTTCTCTCTCTCTCAAAATGTCTCATTTCTTCAACCA	3348
Db	28477	TTTGTTTTCCCTCTGCGCTCTTTTCTCTCTCTCTCAAAATGTCTCATTTCTTCAACCA	2841
QY	3349	GTTCCCTAATGTTCCTGGGGATGCATCTGACCTTTCATATACCTTCCCTCAGTATC	3408
Db	28417	GTTCCCTAATGTTCCTGGGGATGCATCTGACCTTTCATATACCTTCCCTCAGTATC	2835
QY	3409	TCAAACATCACTTGAGCTCTGAGGAATATCTATGCTGTGACACTGTGATCTAATCTACT	3468
Db	28357	TCAAACATCACTTGAGCTCTGAGGAATATCTATGCTGTGACACTGTGATCTAATCTACT	2829
QY	3469	TTTCTGAGTCTCAGACATCTCTTTTCAATTTGATGTTCTTACAGGACCTTAAATTCAGAT	3528
Db	28297	TTTCTGAGTCTCAGACATCTCTTTTCAATTTGATGTTCTTACAGGACCTTAAATTCAGAT	2823
QY	3529	CCCCCAACTTAAGCTTTGACATCTTCTTAAACAACACTTTCCTCTGTGTTCTGTT	3588
Db	28237	CCCCCAACTTAAGCTTTGACATCTTCTTAAACAACACTTTCCTCTGTGTTCTGTT	2817
QY	3589	TCAGTAAATGACCCCAAAATGTGCTGATTACTACAAACCAAGTGACAACAAGGCTTCAT	3648
Db	28177	TCAGTAAATGACCCCAAAATGTGCTGATTACTACAAACCAAGTGACAACAAGGCTTCAT	2811
QY	3649	GATCTGGGCTTGGTTATCTTCTAGGTTTATCTCTCCCTGACATTCACGTGTGTC	3708

/rpt family="L2"
9656. .9676
repeat_region /rpt family="TTCA)n"
complement(9677. .10005)
repeat_region /rpt family="L2"
10006. .10186
repeat_region /rpt family="MSTP"
10187. .10496
repeat_region /rpt family="Alusq"
10497. .10671
repeat_region /rpt family="MSTP"
complement(10672. .10981)
repeat_region /rpt family="Aluy"
complement(11053. .11484)
repeat_region /rpt family="L2"
complement(12337. .12474)
repeat_region /rpt family="FLAM C"
complement(12507. .12954)
repeat_region /rpt family="L2"
12956. .13714
repeat_region /rpt family="LIMA7"
complement(14073. .14372)
repeat_region /rpt family="Alusg"
14466. .14974
repeat_region /rpt family="MLTK"
complement(14975. .15266)
repeat_region /rpt family="Alusx"
15267. .15301
repeat_region /rpt family="MLTK"
15337. .15440
repeat_region /rpt family="LTR50"
15635. .15817
repeat_region /rpt family="LTR33A"
complement(15824. .15945)
repeat_region /rpt family="MIR"
16067. .16090
repeat_region /rpt family="AT rich"
complement(16091. .16371)
repeat_region /rpt family="AluBo"

Query Match 99.7%; Score 9973.2; DB 9; Length 112295;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 9997; Conservative 0; Mismatches 3; Indels 12; Gaps 1;

1 TCCTCTCCCAAGGCGCCAGCAAAATCTGTAGATTCAACAGAGTTCTGACAGCTG 60
31777 TCCTCTCCCAAGGCGCCAGCAAAATCTGTAGATTCAACAGAGTTCTGACAGCTG 31718
61 AGACAAGTTTGTGAGAAATCTGTATGAGATCATGGGGTCTCAGAGGAGAAATA 120
31717 AGACAAGTTTGTGAGAAATCTGTATGAGATCATGGGGTCTCAGAGGAGAAATA 31658
121 TAAAGTTTCAAGAGGTGAGAGGAAAGAAAGGTGAGGGAGTCTTAGAATAGTGGCTC 180
31657 TAAAGTTTCAAGAGGTGAGAGGAAAGAAAGGTGAGGGAGTCTTAGAATAGTGGCTC 31598
181 CCATTGCCCAACACCCAGAAAGACATGCCCTGTCAATGGGAGAGAGGTAGATAGAA 240
31597 CCATTGCCCAACACCCAGAAAGACATGCCCTGTCAATGGGAGAGAGGTAGATAGAA 31538
241 CATTGGCTGTACAGCGATGCGATTGCCAGGCTGCGAAGACTCAGAGATGCCCTT 300
31537 CATTGGCTGTACAGCGATGCGATTGCCAGGCTGCGAAGACTCAGAGATGCCCTT 31478
301 GCCCACTGACTATGAGAGGAGATGATGTTCAACAGCAATTTTCTTGTAGTACAGA 360
31477 GCCCACTGACTATGAGAGGAGATGATGTTCAACAGCAATTTTCTTGTAGTACAGA 31418
361 GAGGACATTTGAGCTGATGCGAGAGGCTGTGTGACATGTTGTTCCAGAGTTCCGAGATG 420
31417 GAGGACATTTGAGCTGATGCGAGAGGCTGTGTGACATGTTGTTCCAGAGTTCCGAGATG 31358

421 TGGTTTCTCTGTGAGAGAACTTCGAGATGAGAAAGGATCTGAGACTTTTGTGA 480
31357 TGGTTTCTCTGTGAGAGAACTTCGAGATGAGAAAGGATCTGAGACTTTTGTGA 31298
481 AGATTATATATGAGACTGTCAAGGGGTCTGAGACCATCTGTAGAGGATCAAGGCCCTTTCA 540
31297 AGATTATATATGAGACTGTCAAGGGGTCTGAGACCATCTGTAGAGGATCAAGGCCCTTTCA 31238
541 GCGTTGGCTAGAGAGAGAGGGGTCTGAGAACTTCACTCTGAGGGGCTGAGCTGAGCCCA 600
31237 GCGTTGGCTAGAGAGAGGGGTCTGAGAACTTCACTCTGAGGGGCTGAGCTGAGCCCA 31178
601 TAATCTTTTCTGACTCACTAGAGCAAAATCTCACAGAAATGAGGAGGAGCTTTGAGAGTGA 660
31177 TAATCTTTTCTGACTCACTAGAGCAAAATCTCACAGAAATGAGGAGGAGCTTTGAGAGTGA 31118
661 GCCCAGAGAGTACTGAGATGAGTACAGGTGAGATCCAGAGAGATGATGTGGGCCGAG 720
31117 GCCCAGAGAGTACTGAGATGAGTACAGGTGAGATCCAGAGAGATGATGTGGGCCGAG 31058
721 ACACTGAGAGAGAAACAGAGACTGTCAATAAAGGCGCTGTGACTCTTAGATCTGAT 780
31057 ACACTGAGAGAGAAACAGAGACTGTCAATAAAGGCGCTGTGACTCTTAGATCTGAT 30998
781 ATGCCCTACACATACCTACACCCCAATCTTAATATTTCTTCACTAGAGGGGGGAG 840
30997 ATGCCCTACACATACCTACACCCCAATCTTAATATTTCTTCACTAGAGGGGGGAG 30938
841 ATGTTCAGAAATTTGGCTGCAACACTAGACACACTAGTAAATGATGATTTTG 900
30937 ATGTTCAGAAATTTGGCTGCAACACTAGACACACTAGTAAATGATGATTTTG 30878
901 CATTCTTTTCACTCAACAAATTTTCTGAAACACTTATATGCGAGCACTATTTTG 960
30877 CATTCTTTTCACTCAACAAATTTTCTGAAACACTTATATGCGAGCACTATTTTG 30818
961 GAGTCAGAGATATATATGATGTAACAGACAGGCAAAACAAAGCAAGCAACAACA 1020
30817 GAGTCAGAGATATATATGATGTAACAGACAGGCAAAACAAAGCAAGCAACAACA 30758
1021 TCACAGATATATAGACATGATGAAAGAAATTTCAAGTTTGTAGTAAATTAACACAC 1080
30757 TCACAGATATATAGACATGATGAAAGAAATTTCAAGTTTGTAGTAAATTAACACAC 30698
1081 AAGGCTGAAATGCTAGATTAAGGGGTCAAGAAAGCTTCAATTGAGAGGTAGCATTT 1140
30697 AAGGCTGAAATGCTAGATTAAGGGGTCAAGAAAGCTTCAATTGAGAGGTAGCATTT 30638
1141 AAGCAGAGTCAAGCTAGAAATTTGTAATTCAGTTACAGTTCAATTTGTTGGGTT 1200
30637 AAGCAGAGTCAAGCTAGAAATTTGTAATTTGTAATTCAGTTCAATTTGTTGGGTT 30578
1201 GGTAAATTAAGCTTTTCCCAAGTGAATCTACCAAGAAAGCTAATTAAGTATAG 1260
30577 GGTAAATTAAGCTTTTCCCAAGTGAATCTACCAAGAAAGCTAATTAAGTATAG 30518
1261 TGGTGTGCTCTCTGAGAGAGACACTCTGTTTCTGCTCATTTACTGCAACCTTTC 1320
30517 TGGTGTGCTCTCTGAGAGAGACACTCTGTTTCTGCTCATTTACTGCAACCTTTC 30458
1321 ACTTCCAGCACTTTTTCGAAAGCCCTTTCAGAGTCAAGGAGGAGGAGGCTGAGCATG 1380
30457 ACTTCCAGCACTTTTTCGAAAGCCCTTTCAGAGTCAAGGAGGAGGAGGCTGAGCATG 30398
1381 GGGCTTGACATTTTGAACAAGTGAACATTTATGCTCCAGACTCATAGGCCCAAGGT 1440
30397 GGGCTTGACATTTTGAACAAGTGAACATTTATGCTCCAGACTCATAGGCCCAAGGT 30338
1441 AAAGCTGAAGAGGCTTGGCATGCCCCAGAAAGGCCCTGATGAAGCTTGAAGAAAGCTG 1500
30337 AAAGCTGAAGAGGCTTGGCATGCCCCAGAAAGGCCCTGATGAAGCTTGAAGAAAGCTG 30278

REFERENCE
AUTHORS

2 (bases 1 to 112295)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bede, F., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
 Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
 Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
 Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Labocque, K.,
 Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Lien, C., Liu, G.,
 Macdonald, P., Margulis, N., McCarthy, M., McEwan, P., McKernan, K.,
 McPeters, R., Meldrim, J., Meneus, L., Mihova, T., Miengo, V.,
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
 O'Donnell, P., O'Neill, D., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
 Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
 Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talama, J., Testaye, S., Theodore, J.,
 Tircelli, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zimmer, A. and Zody, M.

TITLE
JOURNAL

Direct Submission
 Submitted (25-OCT-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (bases 1 to 112295)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, E., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R.,
 Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
 Mihova, T., Miengo, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phukhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talama, J., Testaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

Direct Submission
 Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

4 (bases 1 to 112295)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, E., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R.,
 Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
 Mihova, T., Miengo, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phukhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talama, J., Testaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

Direct Submission
 Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jun 1, 2002 this sequence version replaced g1:19683126.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L10956

Center clone name: 2010_I_16

Only the last 112.3 kb of this clone are being submitted.
 The remainder overlaps AC068733 (WICGR project L10266).

FEATURES

source

1. 112295
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="CTD-2010116"
 /clone_1lb="CTDI Human BAC"
 complement(949..1264)
 /rpt_family="AluSg"
 repeat_region
 1269..1568
 /rpt_family="L2"
 complement(2334..2584)
 /rpt_family="AluSg"
 complement(2429..2433)
 /note="<30 qual SINGL region"
 complement(2462..2466)
 /note="<30 qual SINGL region"
 complement(2478..2503)
 /note="<30 qual SINGL region"
 complement(2508..2537)
 /note="<30 qual SINGL region"
 complement(2577..2583)
 /note="<30 qual SINGL region"
 repeat_region
 2772..3185
 /rpt_family="L2"
 complement(4149..4288)
 /rpt_family="L2"
 repeat_region
 4268..4343
 /rpt_family="L2"
 repeat_region
 4453..4524
 /rpt_family="L2"
 repeat_region
 4563..5531
 /rpt_family="L2"
 complement(5560..5857)
 /rpt_family="AluSx"
 5861..6365
 /rpt_family="L2"
 complement(6378..7246)
 /rpt_family="L2"
 repeat_region
 7898..7942
 /rpt_family="GC_rich"
 repeat_region
 8243..8264
 /rpt_family="GC_rich"
 complement(8355..8589)
 /rpt_family="MIR"
 complement(8875..9148)
 /rpt_family="AluJo"
 9264..9349
 /rpt_family="L2"
 complement(9361..9480)
 /rpt_family="L1MC5"
 complement(9584..9639)
 repeat_region

Db 173311 ACAACTTGGCTTGACGTGACCATCATG-TCCTCCACTTTCCTCACTACTTCATGATC 173253
 Qy 8101 TTTCAGTCTCAGTTCACATGATACCTCCCTCAGTTCCTCTTTCTAGTAAGATTC 8160
 Db 173252 TTTCAGTCTCAGTTCACATGATACCTCCCTCAGTTCCTCTTTCTAGTAAGATTC 173193
 Qy 8161 CAGAGAGGAAATCTGAATGGCCCAAGTCATATTTTTCAGACACACACACATTAAGTGGT 8220
 Db 173192 CAGAGAGGAAATCTGAATGGCCCAAGTCATATTTTTCAGACACACACATTAAGTGGT 173133
 Qy 8221 GATTCGACCTCTATGATTTGGCTACATTAATGGGTTGGAACTCATCTTACTTCAATG 8280
 Db 173132 GATTCGACCTCTATGATTTGGCTACATTAATGGGTTGGAACTCATCTTACTTCAATG 173073
 Qy 8281 CACAAAGACGATAGTCTGGTCTCAAAATAGAGCCCTGGGCGAGGTGGTGGCTCA 8340
 Db 173072 CACAAAGACGATAGTCTGGTCTCAAAATAGAGCCCTGGGCGAGGTGGTGGCTCA 173013
 Qy 8341 TGCCTATTAATCCCAACACTGTGGAGCCGAGGGGGGAGATCATCTTGAATCCAGAGTT 8400
 Db 173012 TGCCTATTAATCCCAACACTGTGGAGCCGAGGGGGGAGATCATCTTGAATCCAGAGTT 172953
 Qy 8401 CTAGACACGCTGGGCAACATGCTGAATCTCATCTCTAATAAATAACAAAATTAAC 8460
 Db 172952 CTAGACACGCTGGGCAACATGCTGAATCTCATCTCTAATAAATAACAAAATTAAC 172893
 Qy 8461 CAGGTGTGGGATGACACAGTATCCAGCTGTTCAGAGGCTGAGGTGGAGGATG 8520
 Db 172892 CAGGTGTGGGATGACACAGTATCCAGCTGTTCAGAGGCTGAGGTGGAGGATG 172833
 Qy 8521 CTGAGTGTGGAGGAGAGATTTGAGTGAACCGTACCTGTGCTGCAATCCAGCTG 8580
 Db 172832 CTGAGTGTGGAGGAGAGATTTGAGTGAACCGTACCTGTGCTGCAATCCAGCTG 172773
 Qy 8581 GGTGACATGTAAGACCTGTCTCAAAAACAAATTAATAAATAATATATGTTTC 8640
 Db 172772 GGTGACATGTAAGACCTGTCTCAAAAACAAATTAATAAATAATATATGTTTC 172713
 Qy 8641 CTGAGAGGGTAAATTCAGTGGGAACTCCAGGGGAGGTGATATGTCAGTCAACCGCT 8700
 Db 172712 CTGAGAGGGTAAATTCAGTGGGAACTCCAGGGGAGGTGATATGTCAGTCAACCGCT 172653
 Qy 8701 GTATACTCAGTACACGGCTAATTAAGAACTTGTGTAGCAGAGAAACATAGTATTT 8760
 Db 172652 GTATACTCAGTACACGGCTAATTAAGAACTTGTGTAGCAGAGAAACATAGTATTT 172593
 Qy 8761 ACTCAACAAATATTTGTTGAGCATCTGATTAAGAGTGGCATTTGCTAGGCACTGAT 8820
 Db 172592 ACTCAACAAATATTTGTTGAGCATCTGATTAAGAGTGGCATTTGCTAGGCACTGAT 172533
 Qy 8821 ACAAGTATGCAATGAGCAGACAAAGATGCTGCCCTGACAGGCTGTCTAAGTAGAGAG 8880
 Db 172532 ACAAGTATGCAATGAGCAGACAAAGATGCTGCCCTGACAGGCTGTCTAAGTAGAGAG 172473
 Qy 8881 GACAAATTAAGAAAGAAAGAAAGAAAGAAATATTTTATAGTATTTTAAAGGTTTAA 8940
 Db 172472 GACAAATTAAGAAAGAAAGAAAGAAAGAAATATTTTATAGTATTTTAAAGGTTTAA 172413
 Qy 8941 GAAATTAAGACAGATAGTGGATTAAGAGTGAAGAGATGAGGCTGTCTTGAAGAA 9000
 Db 172412 GAAATTAAGACAGATAGTGGATTAAGAGTGAAGAGATGAGGCTGTCTTGAAGAA 172353
 Qy 9001 TGAATTTTGAAGCTTCACTTCACTGTAAGAGAAATTAACACACAGATGCTGAGAA 9060
 Db 172352 TGAATTTTGAAGCTTCACTTCACTGTAAGAGAAATTAACACACAGATGCTGAGAA 172293
 Qy 9061 AAAGCAATTTAGAGAGGCTGAGACAGACATCTTCAAGAAATCAAGAGAAAGCTGTG 9120
 Db 172292 AAAGCAATTTAGAGAGGCTGAGACAGACATCTTCAAGAAATCAAGAGAAAGCTGTG 172233
 Qy 9121 AGGCTTGAACACAGAGAAAGACAGGTGGGTGCTTGAAGAGGACAGGACGCTGAGCC 9180
 Db 172232 AGGCTTGAACACAGAGAAAGACAGGTGGGTGCTTGAAGAGGACAGGACGCTGAGCC 172173

Qy 9181 AGGTACTAGACCTGGTAAGGGTTTCAACCAATAAAGGAGTCATCAGAAAGCTTGA 9240
 Db 172172 AGGTACTAGACCTGGTAAGGGTTTCAACCAATAAAGGAGTCATCAGAAAGCTTGA 172113
 Qy 9241 GCAGGGCTGTGATATATTTTCACTCAATTTTATTAAGATCACTGACTTTTTCAG 9300
 Db 172112 GCAGGGCTGTGATATATTTTCACTCAATTTTATTAAGATCACTGACTTTTTCAG 172053
 Qy 9301 ACATAGTATTAAGATTAACACATGTAAGAGAAATCCAGTACCAATCCGTGAGTT 9360
 Db 172052 ACATAGTATTAAGATTAACACATGTAAGAGAAATCCAGTACCAATCCGTGAGTT 171993
 Qy 9361 GTCCAAATTAAGAGTATGACCGCTTGAGATGATAGACAGAGAGGTGGAGAA 9420
 Db 171992 GTCCAAATTAAGAGTATGACCGCTTGAGATGATAGACAGAGAGGTGGAGAA 171933
 Qy 9421 TCACCATATATATTTTGAAGGTAGAGCTGACAGCATTAATAGTAAATAGAGCCG 9480
 Db 171932 TCACCATATATATTTTGAAGGTAGAGCTGACAGCATTAATAGTAAATAGAGCCG 171873
 Qy 9481 GGTGTGTGGCTTAAAGCTGTATCTTACACATTTTGGAGGCCAAGGCGATGATCAC 9540
 Db 171872 GGTGTGTGGCTTAAAGCTGTATCTTACACATTTTGGAGGCCAAGGCGATGATCAC 171813
 Qy 9541 TGAGGTGAGAGTGAAGACAGCTTGACCAACATGTAACCTGTCTTACTTAATAA 9600
 Db 171812 TGAGGTGAGAGTGAAGACAGCTTGACCAACATGTAACCTGTCTTACTTAATAA 171753
 Qy 9601 TACAAATTAAGTGGAAATGTTGGACATGCTGTATCTACGCTTACTTGGAGGCTGA 9660
 Db 171752 TACAAATTAAGTGGAAATGTTGGACATGCTGTATCTACGCTTACTTGGAGGCTGA 171693
 Qy 9661 GGCAGAGAAATGCTTGAACCTGGAGGTGAATGTTGACATGAGCCGAAATTCACACT 9720
 Db 171692 GGCAGAGAAATGCTTGAACCTGGAGGTGAATGTTGACATGAGCCGAAATTCACACT 171633
 Qy 9721 GCACTCCAGCTGGGGAAACAGAGTGAACCTCGCTCTTAATTAATGAATGAATG 9780
 Db 171632 GCACTCCAGCTGGGGAAACAGAGTGAACCTCGCTCTTAATTAATGAATGAATG 171573
 Qy 9781 ATATCAGTCAAGTGAAGGAAAGGAAAGGCTTCAAGATGACTCAGCTTTCGTGACT 9840
 Db 171572 ATATCAGTCAAGTGAAGGAAAGGAAAGGCTTCAAGATGACTCAGCTTTCGTGACT 171513
 Qy 9841 CAGCACTGAGTGGCTGTGTTTCTTAAATTTGGAAAGCTAGGAGTGTG 9900
 Db 171512 CAGCACTGAGTGGCTGTGTTTCTTAAATTTGGAAAGCTAGGAGTGTG 171453
 Qy 9901 TGTGTGGTGGGGGCAAAATCAGTTTGGCATTAATAGTTTGGGCTATTTGGCACCC 9960
 Db 171452 TGTGTGGTGGGGGCAAAATCAGTTTGGCATTAATAGTTTGGGCTATTTGGCACCC 171393
 Qy 9961 CATTAAGCATGTCAGGTAGGACGCTGATTTGAGCCTTAAC 10000
 Db 171392 CATTAAGCATGTCAGGTAGGACGCTGATTTGAGCCTTAAC 171353

RESULT 2
 AC084337/c 112295 bp DNA linear pri 01-JUN-2002
 LOCUS AC084337 Homo sapiens chromosome 11, clone CTD-2010116, complete sequence.
 DEFINITION AC084337
 ACCESSION AC084337.7 GI:21306722
 VERSION AC084337.7
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 112295)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 11, clone CTD-2010116
 JOURNAL Unpublished

Db 177691 GGTATCTCTCAGGTTTATCTCTCCCTCCGCAATTCAGTCTGTGCGCAGCATAGCAA 177632
Qy 3721 TCTACATGAGGTTGAGAGCACTGCTTCCATAGTTTGGGCTGTGATGTGCTCCCTCT 3780
Db 177631 TCTACATGAGGTTGAGAGCACTGCTTCCATAGTTTGGGCTGTGATGTGCTCCCTCT 177572
Qy 3781 GCTGTGACACCCCTTCTCTCACTTGTCAACCTTGAAAATTCCTGTGATTTTCACTCT 3840
Db 177571 GCTGTGACACCCCTTCTCTCACTTGTCAACCTTGAAAATTCCTGTGATTTTCACTCT 177512
Qy 3841 TGGGCCCAATGCTCTCTTGTGTGAAACCTTCCAACTTCTCTAGGCACTTAAAG 3900
Db 177511 TGGGCCCAATGCTCTCTTGTGTGAAACCTTCCAACTTCTCTAGGCACTTAAAG 177452
Qy 3901 CACTGTCTATATCTCAGTGTGACTTTCACCTTACCTTGGTGTGAGGCTAGG 3960
Db 177451 CACTGTCTATATCTCAGTGTGACTTTCACCTTACCTTGGTGTGAGGCTAGG 177392
Qy 3961 ATTGCAAGAGTCTTCTCTGCTTGTGACAGTGAACCTTCTGAAGTGAAGAAGAGAGTCTT 4020
Db 177391 ATTGCAAGAGTCTTCTCTGCTTGTGACAGTGAACCTTCTGAAGTGAAGAAGAGAGTCTT 177332
Qy 4021 GTTATCTCAGTGTCTCTCAACATGCTGCGCATATAGTATTAATGAGTACTGTTCTT 4080
Db 177331 GTTATCTCAGTGTCTCTCAACATGCTGCGCATATAGTATTAATGAGTACTGTTCTT 177272
Qy 4081 GGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 4140
Db 177271 GGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 177212
Qy 4141 ATTGCAAGATGAACAGTGTGTGATGTTTGTCAACAGTGTGATGTGTGAGATTAATGT 4200
Db 177211 ATTGCAAGATGAACAGTGTGTGATGTTTGTCAACAGTGTGATGTGTGAGATTAATGT 177152
Qy 4201 GCCACAGAGTGTCTGGGTACAGAGCTAGAGCATGTGTATTAATGATGACTGGA 4260
Db 177151 GCCACAGAGTGTCTGGGTACAGAGCTAGAGCATGTGTATTAATGATGACTGGA 177092
Qy 4261 TTTGCAAACTGAGAGTGTGATTAATGCAAAAGAGCAGACATTTTGTGCCAATGTG 4320
Db 177091 TTTGCAAACTGAGAGTGTGATTAATGCAAAAGAGCAGACATTTTGTGCCAATGTG 177032
Qy 4321 ACTGAGAAATGTGAGGGCCACAGAAAGATATGATTAAGCAGTATGATTAATGTGTG 4380
Db 177031 ACTGAGAAATGTGAGGGCCACAGAGATATGATTAAGCAGTATGATTAATGTGTG 176972
Qy 4381 GTTAATGCAAGTGTGAGTATCTGGGATGCAAGTCAAAAAGAGATCTTTGAATGC 4440
Db 176971 GTTAATGCAAGTGTGAGTATCTGGGATGCAAGTCAAAAAGAGATCTTTGAATGC 176912
Qy 4441 AGGGGACAAAGTCTGGGTATACCTCTCGTGAAGAAAGAGAAAGATACCCAAATGTGC 4500
Db 176911 AGGGGACAAAGTCTGGGTATACCTCTCGTGAAGAAAGAGAAAGATACCCAAATGTGC 176852
Qy 4501 TCCAAAGATGAATTTCTGGAATCCCATCCCATGTGATGAGTGTGATGATGATGATGATG 4560
Db 176851 TCCAAAGATGAATTTCTGGAATCCCATCCCATGTGATGAGTGTGATGATGATGATG 176792
Qy 4561 GAGATGTCAAGGTGAAGGCTCTCTTCTTCCAGAGTCAAGTCCAGGCTGAATCCAGAG 4620
Db 176791 GAGATGTCAAGGTGAAGGCTCTCTTCTTCCAGAGTCAAGTCCAGGCTGAATCCAGAG 176732
Qy 4621 ACCGTGAGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4680
Db 176731 ACCGTGAGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 176672
Qy 4681 GTGTGTGTGAGGCACTATGTGGCTTTTAAAGGAAGGAATGAAGAGCAAGTATGTGTGT 4740
Db 176671 GTGTGTGTGAGGCACTATGTGGCTTTTAAAGGAAGGAATGAAGAGCAAGTATGTGTGT 176612
Qy 4741 CTAGAGTCACTGAGGGAAGAGAGAGTCAATCTGTGTGAAGGCACTGTGTGTATTT 4800
Db 176611 CTAGAGTCACTGAGGGAAGAGAGAGTCAATCTGTGTGAAGGCACTGTGTGTATTT 176552

Qy 4801 CTGTGTGAATGAGTCCCTCAATGTTTCCCATGAGCCCTTGAAGCAATCTACCTCTTATGT 4860
Db 176551 CTGTGTGAATGAGTCCCTCAATGTTTCCCATGAGCCCTTGAAGCAATCTACCTCTTATGT 176492
Qy 4861 CTTCATGAAGTCTCTCAATGAGCAAAATCCCTATTTGGCCCTTCACTTGTATTTATTT 4920
Db 176491 CTTCATGAAGTCTCTCAATGAGCAAAATCCCTATTTGGCCCTTCACTTGTATTTATTT 176432
Qy 4921 ATACTTTACCTTAAACATGAAGTCCAGAAACCCATATGCTATCTGTATCACTGATTT 4980
Db 176431 ATACTTTACCTTAAACATGAAGTCCAGAAACCCATATGCTATCTGTATCACTGATTT 176372
Qy 4981 GCTCAAGTCTTGAATCTTCTCTCTGAGGGGTCCATCTCTCTGAGTCTTTTA 5040
Db 176371 GCTCAAGTCTTGAATCTTCTCTCTGAGGGGTCCATCTCTCTGAGTCTTTTA 176312
Qy 5041 ATTCTTACTTCAAGATTTCACTTAAATATCATCTTCCCTGAGAAATTTTCCAGACTCTC 5100
Db 176311 ATTCTTACTTCAAGATTTCACTTAAATATCATCTTCCCTGAGAAATTTTCCAGACTCTC 176252
Qy 5101 CCCACTGCTTGTGAGTGAATCTGTGTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5160
Db 176251 CCCACTGCTTGTGAGTGAATCTGTGTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 176192
Qy 5161 CCGTCTTTAGGCATCTCTGATGAGCTGAGCTCATAGTGTGAGTCAATATCTGG 5220
Db 176191 CCGTCTTTAGGCATCTCTGATGAGCTGAGCTCATAGTGTGAGTCAATATCTGG 176132
Qy 5221 CCCATCTGAGAACCCAGAGAAAGCAAGAGAGGGGTAAACCGGTCTCAACCAATGCTCT 5280
Db 176131 CCCATCTGAGAACCCAGAGAAAGCAAGAGAGGGGTAAACCGGTCTCAACCAATGCTCT 176072
Qy 5281 GTTATGATTTGAACAAAGTGAACCGGAGTGTCTTGGGACTTGGCTACGGGAAACATG 5340
Db 176071 GTTATGATTTGAACAAAGTGAACCGGAGTGTCTTGGGACTTGGCTACGGGAAACATG 176012
Qy 5341 AAGGAGCGTTCCTGAGCAGTGTGGGAACTGCTCTGCTGCTGCTGAGATGAGTGGCCG 5400
Db 176011 AAGGAGCGTTCCTGAGCAGTGTGGGAACTGCTCTGCTGCTGAGATGAGTGGCCG 175952
Qy 5401 TACTACTGCTTCAGAGGTAAACCAATTCCTGCGCTTGAACCTGTCAAGGGAGAGTGTCT 5460
Db 175951 TACTACTGCTTCAGAGGTAAACCAATTCCTGCGCTTGAACCTGTCAAGGGAGAGTGTCT 175892
Qy 5461 CCCAGGTACCCGCGGAGTGTCCGAGACTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520
Db 175891 CCCAGGTACCCGCGGAGTGTCCGAGACTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175832
Qy 5521 GCCCTAGCACTTGAACCTGTGAGATTCATCATCTTCCCTGAGCTGTGATCTCACG 5580
Db 175831 GCCCTAGCACTTGAACCTGTGAGATTCATCATCTTCCCTGAGCTGTGATCTCACG 175772
Qy 5581 TGTCTAGTCTCACTTAACTCCGTGTGCAACCTTGTGAGCTTAACTAGAGCCCATTT 5640
Db 175771 TGTCTAGTCTCACTTAACTCCGTGTGCAACCTTGTGAGCTTAACTAGAGCCCATTT 175712
Qy 5641 TCCATCTGGAATTTCCATTCCTGCTCATATGAGGAAACCCACACCCCACTAACCCACAG 175652
Db 175711 TCCATCTGGAATTTCCATTCCTGCTCATATGAGGAAACCCACACCCCACTAACCCACAG 175600
Qy 5701 CATCTCTTCACTTGAACCTGATCTGATCTGTGCTCTGTGCTGTGTGTGTGTGTGTGTGTGT 5760
Db 175651 CATCTCTTCACTTGAACCTGATCTGATCTGTGCTCTGTGCTGTGTGTGTGTGTGTGTGTGT 175592
Qy 5761 ATTTCTCTCAAGGCACTGAGCAACAGAAATGAGATGAGGATGAGGATGAGGATGAGGATGAG 5820
Db 175591 ATTTCTCTCTCAAGGCACTGAGCAACAGAAATGAGATGAGGATGAGGATGAGGATGAGGATGAG 175532
Qy 5821 TGGCCCTGAGTATGAGGCTGTGAGCCCACTATGATCTTGTGTGCACTGACGTCTGACAA 5880
Db 175531 TGGCCCTGAGTATGAGGCTGTGAGCCCACTATGATCTTGTGTGCACTGACGTCTGACAA 175472

```
QY 1501 TTCTGAGATTTCTAAGTAAGTTATCTGNGTGTGTTACTAAAGTAAAGTAT 1560
DB 179851 TTCTGAGATTTCTAAGTAAGTTATCTGNGTGTGTTACTAAAGTAAAGTAT 179792
QY 1561 TGTGTCTCTAGTGTGCTTATAGACAGAGGCTTGAACAAGTACACAGCAATATTAGTCCCT 1620
DB 179791 TGTGTCTCTAGTGTGCTTATAGACAGAGGCTTGAACAAGTACACAGCAATATTAGTCCCT 179732
QY 1621 CCTTTTCTACCTCCCTCATTTGTGAGATTAATCTCAATCAACAAAGGTATCTCACTCT 1680
DB 179731 CCTTTTCTACCTCCCTCATTTGTGAGATTAATCTCAATCAACAAAGGTATCTCACTCT 179672
QY 1681 ACTCACTCCCTGACTTATAGATGCTGAGCCATTTGCCAGTGTGAGTCAAGCTGGA 1740
DB 179671 ACTCACTCCCTGACTTATAGATGCTGAGCCATTTGCCAGTGTGAGTCAAGCTGGA 179612
QY 1741 CGTCAAGTGTAGCCAGCTTACTGTCTGTAATAATGCTGAAGGGGTTGGGGGCAAGCTG 1800
DB 179611 CGTCAAGTGTAGCCAGCTTACTGTCTGTAATAATGCTGAAGGGGTTGGGGGCAAGCTG 179552
QY 1801 CCGGGAAAAAGAGTCTTGTGATTCAGATTTCTGTCCAGACCTTGACCTTATTGCACTGA 1860
DB 179551 CCGGGAAAAAGAGTCTTGTGATTCAGATTTCTGTCCAGACCTTGACCTTATTGCACTGA 179492
QY 1861 TGTATCAGCAATATTGCTGTAGTCTGCTGAGAGACAGCAATCCAGTATGATGGAAG 1920
DB 179491 TGTATCAGCAATATTGCTGTAGTCTGCTGAGAGACAGCAATCCAGTATGATGGAAG 179432
QY 1921 TGGGGGTGTGTCTGTCTGCAACTCTATATAGGAGATTCACTGTGCAACCCAGAGCTGTCC 1980
DB 179431 TGGGGGTGTGTCTGTCTGCAACTCTATATAGGAGATTCACTGTGCAACCCAGAGCTGTCC 179372
QY 1981 TGTGGCTCTGTGACGTCTCAGATGAGTCTAGGTTACTGGAGACACCCGTTGCACTGGGTTGT 2040
DB 179371 TGTGGCTCTGTGACGTCTCAGATGAGTCTAGGTTACTGGAGACACCCGTTGCACTGGGTTGT 179312
QY 2041 GGAGCCTATGCTGTCTGTGCAATGCGCACCCCTCTCTCGGTGATGATGAGTGGAG 2100
DB 179311 GGAGCCTATGCTGTCTGTGCAATGCGCACCCCTCTCTCGGTGATGATGAGTGGAG 179252
QY 2101 TGAAGCGAAGATTGATTTCTGGGCTAGGGTAAAGTAAAGGCGCAAGTTTAAAGCCTCGGT 2160
DB 179251 TGAAGCGAAGATTGATTTCTGGGCTAGGGTAAAGTAAAGGCGCAAGTTTAAAGCCTCGGT 179192
QY 2161 CAAATTTGGGGTCAAGGGGCTATAGGAAAGGAGTGGTCCCAATGATCAAGATATCTATT 2220
DB 179191 CAAATTTGGGGTCAAGGGGCTATAGGAAAGGAGTGGTCCCAATGATCAAGATATCTATT 179132
QY 2221 TTGTTCTCCCTAGGACTAGTGGCCATGAGGAATGTGCTGAAGGCGAGACCAAGCCAGACC 2280
DB 179131 TTGTTCTCCCTAGGACTAGTGGCCATGAGGAATGTGCTGAAGGCGAGACCAAGCCAGACC 179072
QY 2281 CAGACGTGACTGTGAGGCTCTGACTCCCTAAGTCTGTCTTATCTGTCTGTGTGTCTC 2340
DB 179071 CAGACGTGACTGTGAGGCTCTGACTCCCTAAGTCTGTCTTATCTGTCTGTGTGTCTC 179012
QY 2341 TGCATTTTATCACTCTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTACTGTTGCACTCTCCT 2400
DB 179011 TGCATTTTATCACTCTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTACTGTTGCACTCTCCT 178952
QY 2401 ACCCTCAACCCAGAAAGCTGCTGAGATGCTGAGGCTTATAGTATCAACCTGAGANTGAC 2460
DB 178951 ACCCTCAACCCAGAAAGCTGCTGAGATGCTGAGGCTTATAGTATCAACCTGAGANTGAC 178892
QY 2461 AATGGAACCATGCTGTTTTTAAAGTAGAGGGAAGCTGAGGCTTAAAGGCTTTAGAACCTT 2520
DB 178891 AATGGAACCATGCTGTTTTTAAAGTAGAGGGAAGCTGAGGCTTAAAGGCTTTAGAACCTT 178832
QY 2521 AGACTTACTCTCTTCAACAAAGGCTCTCTGTCTGTGGAGAGCTTAAAGAAATATCTGA 2580
DB 178831 AGACTTACTCTCTTCAACAAAGGCTCTCTGTCTGTGGAGAGCTTAAAGAAATATCTGA 178772
QY 2581 TGGATCACTGACAGCTTCTCTCAAGCTATCTCAGTAGAGTCAAAAGTTTCTCACTGGGCC 2640
DB 178771 TGGATCACTGACAGCTTCTCTCAAGCTATCTCAGTAGAGTCAAAAGTTTCTCACTGGGCC 178712
QY 2641 CCTCAGTGAAGTGGTTTTTCAAGGAGATTGTGTGTGAAGATCAAAATGGGACCCG 2700
DB 178711 CCTCAGTGAAGTGGTTTTTCAAGGAGATTGTGTGTGAAGATCAAAATGGGACCCG 178652
QY 2701 GAGTTAATCTCAGAGAGATGGAAGAAATTTCCCAAGCCGTGTGATGCTCAATCCGTCAA 2760
DB 178651 GAGTTAATCTCAGAGAGATGGAAGAAATTTCCCAAGCCGTGTGATGCTCAATCCGTCAA 178592
QY 2761 GGTCAACAGTGTCTTTCTGATCAAGTACTGTGTGGCCAAATATGAGGCGCAGCTGGA 2820
DB 178591 GGTCAACAGTGTCTTTCTGATCAAGTACTGTGTGGCCAAATATGAGGCGCAGCTGGA 178532
QY 2821 AAGGGCTGGAATCGACACTGGGGACCTTTCCCAATGAGCTTGTGCAATGAGCCCATAG 2880
DB 178531 AAGGGCTGGAATCGACACTGGGGACCTTTCCCAATGAGCTTGTGCAATGAGCCCATAG 178472
QY 2881 CAATAGTAGAGATTTCTTTCCCATGAGCCCTCTCTGTGTAAGGCTGGGCTAAGG 2940
DB 178471 CAATAGTAGAGATTTCTTTCCCATGAGCCCTCTCTGTGTAAGGCTGGGCTAAGG 178412
QY 2941 GAGTGTGATGCTGTGGGCTGTGGCAGTGTGCAACATCGAGTGTGTTCTTCACTAG 3000
DB 178411 GAGTGTGATGCTGTGTGGGCTGTGGCAGTGTGCAACATCGAGTGTGTTCTTCACTAG 178352
QY 3001 TCTTAGTTTCAACCGCTCTGCTGTAACCTCAACTGTGAGCCATCTTTTCTCTGAGC 3060
DB 178351 TCTTAGTTTCAACCGCTCTGCTGTAACCTCAACTGTGAGCCATCTTTTCTCTGAGC 178292
QY 3061 AATTGCTTCCCTGCTTCCATGACCCCTGATCAATCTCTTCAATAGGCAAGGCAAGT 3120
DB 178291 AATTGCTTCCCTGCTTCCATGACCCCTGATCAATCTCTTCAATAGGCAAGGCAAGT 178232
QY 3121 TGTTCACCAACAATGGACCTGTGATGAGAGACATGTGGAGCATGAAATCCAGTGTGCT 3180
DB 178231 TGTTCACCAACAATGGACCTGTGATGAGAGACATGTGGAGCATGAAATCCAGTGTGCT 178172
QY 3181 GTGTCTCAACAGTCCCATGTGACCCAGGCTGTGTCTGTCTCAGAGAAAGGGGTCTTTTC 3240
DB 178171 GTGTCTCAACAGTCCCATGTGACCCAGGCTGTGTCTGTCTCAGAGAAAGGGGTCTTTTC 178112
QY 3241 CTATCTTGTGCAAAAGTGTGTGTGTGTGGGAAAGTCTGATCGGCTTGTGTTCCCT 3300
DB 178111 CTATCTTGTGCAAAAGTGTGTGTGTGTGGGAAAGTCTGATCGGCTTGTGTTCCCT 178052
QY 3301 CTTGCTCTTTTCTCTCTCTCTCAATGTCTCATTTCTCAATCTCAACCAAGTCCCTAATGT 3360
DB 178051 CTTGCTCTTTTCTCTCTCTCTCAATGTCTCATTTCTCAATCTCAACCAAGTCCCTAATGT 177992
QY 3361 TCTTGGGGAATCATCTAGACCTTTTCCATATACCTTTCCCTCAGTATCTCAACCATAC 3420
DB 177991 TCTTGGGGAATCATCTAGACCTTTTCCATATACCTTTCCCTCAGTATCTCAACCATAC 177932
QY 3421 TTGGCTCTGAGAAATATATGTGTGTGGAACACTGATATAGAACTTCTTGAAGCTCCA 3480
DB 177931 TTGGCTCTGAGAAATATATGTGTGTGTGGAACACTGATATAGAACTTCTTGAAGCTCCA 177872
QY 3481 GACATCTCTTTCAAATGATGTTCTATAGAGGCACTAAATATGAGATCCCCAAACTAA 3540
DB 177871 GACATCTCTTTCAAATGATGTTCTATAGAGGCACTAAATATGAGATCCCCAAACTAA 177812
QY 3541 GCTTTCATCTTCTTAAACCAACCTTCTCTGTGTGTTCTCTGTTCAAGTAAATGAC 3600
DB 177811 GCTTTCATCTTCTTAAACCAACCTTCTCTGTGTGTTCTCTGTTCAAGTAAATGAC 177752
QY 3601 CCAAAATGTGCTGATTAATAAACAAGTCAACAGGCTGATGATCTGGGCTT 3660
DB 177751 CCAAAATGTGCTGATTAATAAACAAGTCAACAGGCTGATGATCTGGGCTT 177692
QY 3661 GGTATCTTCTCAGGTTTATCTCTCCCTGCAATTCAGTGTGTGCCAGCATAGAA 3720
DB 177691 GGTATCTTCTCAGGTTTATCTCTCCCTGCAATTCAGTGTGTGCCAGCATAGAA 177632
```

```

unsure /note="<30 qual SNGL region"
14582. .14588
/note="<30 qual SNGL region"
unsure 14899. 14904
/note="<30 qual SNGL region"
repeat_region 16077. 16386
/rpc_family="Alusx"
repeat_region 16387. 16988
/rpc_family="L1MA3"
complement(17002. .17075)
/rpc_family="MIR"
repeat_region complement(17907. .18207)
/rpc_family="Alusx"
repeat_region complement(18218. .18368)
/rpc_family="MIR"
repeat_region 18430. 18711
/rpc_family="Alusx"
repeat_region 18712. 18743
/rpc_family="AT_rich"
repeat_region 19184. 19685
/rpc_family="L2"
complement(19738. .19892)
/rpc_family="L2"
repeat_region 19894. 20155
/rpc_family="Alusx"
repeat_region 20167. 20189
/rpc_family="(CAAA)n"
complement(20649. 20961)
repeat_region 21292. 21411
/rpc_family="Alusq"
repeat_region 21454. 21594
/rpc_family="L2"
repeat_region 21774. 22026
/rpc_family="MLTIE"
repeat_region 22048. 22570
/rpc_family="MER41A"
repeat_region 22571. 23617
/rpc_family="MER41A-int"
repeat_region 23618. 23651
/rpc_family="(TG)n"

Query Match 99.8%; Score 9983.2; DB 9; Length 191656;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9996; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 TCCCTCTCCCGAGGCCAGCCAGCAAAATCTGTAGATTCAAGAGGTTCTGACAGCTG 60
DB 181351 TCCCTCTCCCGAGGCCAGGCCAGCAAAATCTGTAGATTCAAGAGGTTCTGACAGCTG 181292
QY 61 AAGACAAGTTGTGAGAAATCTGTAGAGGATCATGCGGTCTCAGAGGAGAAATA 120
DB 181291 AAGACAAGTTGTGAGAAATCTGTAGAGGATCATGCGGTCTCAGAGGAGAAATA 181232
QY 121 TAAGTTTCAAGGCTGAGAGGAGAAAGAGTGAAGGGAGTCTTAGAATAGTGGCTC 180
DB 181231 TAAGTTTCAAGGCTGAGAGGAGAAAGAGTGAAGGGAGTCTTAGAATAGTGGCTC 181172
QY 181 CCATTGCCCAACACCCAGAAAGAACATGCTCCTGCAATGGGAGAAAGGTAGATGAGA 240
DB 181171 CCATTGCCCAACACCCAGAAAGAACATGCTCCTGCAATGGGAGAAAGGTAGATGAGA 181112
QY 241 CATTGCTGTAGAGGATGAGCATTTGCCAGGCTGCAAGAGATTCAGAGAGTCCACCTT 300
DB 181111 CATTGCTGTAGAGGATGAGCATTTGCCAGGCTGCAAGAGATTCAGAGAGTCCACCTT 181052
QY 301 GCCCACTGACTATGAGAGGAGAAATGATGTTCAAGACATTTTCATTGTAAGTCAGA 360
DB 181051 GCCCACTGACTATGAGAGGAGAAATGATGTTCAAGACATTTTCATTGTAAGTCAGA 180992
QY 361 GAGGACATTGAGCTGATGAGCAGAGGCTGTGTGACATGTTTTCAGAGGTTCCGAAATG 420
DB 180991 GAGGACATTGAGCTGATGAGCAGAGGCTGTGTGACATGTTTTCAGAGGTTCCGAAATG 180932

```

```

QY 421 TGTGTTTCTGTGGAAGAACTTCGAGAGTAGAAAAAGGATCTGAGACTTTTGTA 480
DB 180931 TGTGTTTCTGTGGAAGAACTTCGAGAGTAGAAAAAGGATCTGAGACTTTTGTA 180872
QY 481 AGATTATATATGGAAGTGTCAAGGGGTCTGAGGCATCTGTAGAGGATCAAGGCCCTTTCA 540
DB 180871 AGATTATATATGGAAGTGTCAAGGGGTCTGAGGCATCTGTAGAGGATCAAGGCCCTTTCA 180812
QY 541 GCTTTGGCTTAGGAGAGAGGGGTCTTGGAATCTTCAATCTGGCCCATAGAGCTGCCCCA 600
DB 180811 GCTTTGGCTTAGGAGAGAGGGGTCTTGGAATCTTCAATCTGGCCCATAGAGCTGCCCCA 180752
QY 601 TAAATCTTTTCTGACTCACTAGGCAAAATCTCAACAGAAATGGGAGGAGCTTTGGAGTGG 660
DB 180751 TAAATCTTTTCTGACTCACTAGGCAAAATCTCAACAGAAATGGGAGGAGCTTTGGAGTGG 180692
QY 661 GCCCAGAAATATCTAGATAGAGATGAGATCCAGAAAGATGATGTGGGCGAG 720
DB 180691 GCCCAGAAATATCTAGATAGAGATGAGATCCAGAAAGATGATGTGGGCGAG 180632
QY 721 ACCTGAGAGAGAAACAGAGCTGTCAATAAAGGCGTCTGTGACTCTTGAATCTCAT 780
DB 180631 ACCTGAGAGAGAAACAGAGCTGTCAATAAAGGCGTCTGTGACTCTTGAATCTCAT 180572
QY 781 ATGCTACTACATACTACCTACCCCAATTCCTAATATTTCTCTACCTAGAGGGGGGAA 840
DB 180571 ATGCTACTACATACTACCTACCCCAATTCCTAATATTTCTCTACCTAGAGGGGGGAA 180512
QY 841 ATTTGAGAAATTTGGCTGCAACACTAGCAACCTACCTAGTAAATGCAATTTTG 900
DB 180511 ATTTGAGAAATTTGGCTGCAACACTAGCAACCTACCTAGTAAATGCAATTTTG 180452
QY 901 CATTTTTTCATTCACAAAATATTTTGGAACAAGCTTATATGCGAGCACTATTTTGA 960
DB 180451 CATTTTTTCATTCACAAAATATTTTGGAACAAGCTTATATGCGAGCACTATTTTGA 180392
QY 961 GAGTCAGGATATATATATGATGTAACAGACAGGCAAAACAAAGCAACACCA 1020
DB 180391 GAGTCAGGATATATATATGATGTAACAGACAGGCAAAACAAAGCAACACCA 180332
QY 1021 TCACGAGATATGATGACAGATGAAAGAAATTCAGTTTATGTAATATAAACAAC 1080
DB 180331 TCACGAGATATGATGACAGATGAAAGAAATTCAGTTTATGTAATATAAACAAC 180272
QY 1081 AAGGCTTGAATATGCTATATAGGCGGTCAAGAAAGCTTCATTGAGAAAGTATGAT 1140
DB 180271 AAGGCTTGAATATGCTATATAGGCGGTCAAGAAAGCTTCATTGAGAAAGTATGAT 180212
QY 1141 AAGCAGAGTCAAGCTAGAAATATGGAATTCAGTTACGTTCAATTTGTTCTGGGTT 1200
DB 180211 AAGCAGAGTCAAGCTAGAAATATGGAATTCAGTTACGTTCAATTTGTTCTGGGTT 180152
QY 1201 GGTAAATATATAGCTTTTCCCAAGGTGAAACTACCAAGAAAGACTAATTAAGTAG 1260
DB 180151 GGTAAATATATAGCTTTTCCCAAGGTGAAACTACCAAGAAAGACTAATTAAGTAG 180092
QY 1261 TGTGTGCTCTCTGGAAGAGACACTCTGTTTCTGCTCATTAAGTCAACCTTTC 1320
DB 180091 TGTGTGCTCTCTGGAAGAGACACTCTGTTTCTGCTCATTAAGTCAACCTTTC 180032
QY 1321 ACTTCCAGGACATTTTTCGAAAGCCCTTGCAAGTCAGGGAAGGAGGCTGGGCATG 1380
DB 180031 ACTTCCAGGACATTTTTCGAAAGCCCTTGCAAGTCAGGGAAGGAGGCTGGGCATG 179972
QY 1381 GGGCTTGAACATTTGCAACAGTGAACATTAATTTGCCAGACTCACTAGCCCAAGGCT 1440
DB 179971 GGGCTTGAACATTTGCAACAGTGAACATTAATTTGCCAGACTCACTAGCCCAAGGCT 179912
QY 1441 AAGCTGAAGAGGCTTTGGCATGCCCCAGAAAGCCCTGATGAAAGCTTGAAGAAAGCTG 1500
DB 179911 AAGCTGAAGAGGCTTTGGCATGCCCCAGAAAGCCCTGATGAAAGCTTGAAGAAAGCTG 179852

```

REFERENCE AUTHORS	TITLE JOURNAL	COMMENT	FEATURES source
2 (bases 1 to 191656) Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G., Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Harford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Kartas, A., Klein, J., Larocque, K., Lamazeres, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Locke, K., Macdonald, P., Margulis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivar, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tessfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	Submitted (08-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. Direct Submission Submitted (12-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Location/Qualifiers 1. 191656 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="11" /map="11" /clone="RP11-304C12" /clone_11b="RPCT-11 Human Male BAC" complement(189..1091) /rpt_family="L2" /rpt_family="GN" complement(1151..1401) /rpt_family="L2" complement(1891..2187) /rpt_family="Aluub" 3141..3165 /rpt_family="CA" complement(3205..3522) /rpt_family="L1MC2" complement(4393..4591) /rpt_family="L2" 4624..4932 /rpt_family="Alus" 4987..5077 complement(5165..6245) /rpt_family="Tigge3b" complement(6246..6555) /rpt_family="Alu" complement(6556..6634) /rpt_family="Tigge3b" 6940..7136 /rpt_family="L2" complement(7607..7849) /rpt_family="L2" complement(11242..11318) /rpt_family="MIR" complement(11385..11572) /rpt_family="MER69A" complement(12325..12520) /rpt_family="MIR" 12536..12706 /rpt_family="MIR" 12905..13529 /rpt_family="L1MA3" 13530..13842 /rpt_family="Alus" 13843..16076 /rpt_family="L1MA3" 14517..14528 /note="<30 qual SNGL region" 14536..14542 /note="<30 qual SNGL region" 14553..14558
3 (bases 1 to 191656) Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kartas, A., Kartas, A., Kells, C., Larocque, K., Lamazeres, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicoll, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhng, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rie, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tessfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	Submitted (27-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Submitted (12-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	repeat_region complement(189..1091) /rpt_family="L2" /rpt_family="GN" complement(1151..1401) /rpt_family="L2" complement(1891..2187) /rpt_family="Aluub" 3141..3165 /rpt_family="CA" complement(3205..3522) /rpt_family="L1MC2" complement(4393..4591) /rpt_family="L2" 4624..4932 /rpt_family="Alus" 4987..5077 complement(5165..6245) /rpt_family="Tigge3b" complement(6246..6555) /rpt_family="Alu" complement(6556..6634) /rpt_family="Tigge3b" 6940..7136 /rpt_family="L2" complement(7607..7849) /rpt_family="L2" complement(11242..11318) /rpt_family="MIR" complement(11385..11572) /rpt_family="MER69A" complement(12325..12520) /rpt_family="MIR" 12536..12706 /rpt_family="MIR" 12905..13529 /rpt_family="L1MA3" 13530..13842 /rpt_family="Alus" 13843..16076 /rpt_family="L1MA3" 14517..14528 /note="<30 qual SNGL region" 14536..14542 /note="<30 qual SNGL region" 14553..14558

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 06:39:26 ; Search time 23513.6 Seconds

(without alignments)
17398.312 Million cell updates/secTitle: US-09-900-448-3_COPY_1_10000
Perfect score: 10000
Sequence: 1 tccctctcccccagcagccgcccc.....agctgatttcgagccttaaac 10000Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl1:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_man:*
37: em_htg_vrc:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	9983.2	99.8	191656	9 AC068733	AC068733 Homo sapi
C 2	9973.2	99.7	112285	9 AC084337	AC084337 Homo sapi
C 3	7556.4	75.6	148198	2 AC024232	AC024232 Homo sapi
C 4	2374.8	23.7	148198	2 AC024232	AC024232 Homo sapi
C 5	1509.4	15.1	5234	6 AX345846	AX345846 Sequence
C 6	1281.6	12.8	5234	6 AX345847	AX345847 Sequence
C 7	681.4	6.8	731	9 HDMXMA01	M36796 Human hemop
C 8	509.4	5.1	511	9 HSEMPEPO	X15399 Human DNA f
C 9	479.6	4.8	249769	2 AC125227	AC125227 Mus muscu
C 10	479.4	4.8	202495	9 CNS01DW6	AL136418 Human chr
C 11	471.6	4.7	74904	9 AC108045	AL139054 Human chr
C 12	467.4	4.7	202496	9 CNS01DX6	AC021105 Human chr
C 13	444.8	4.4	187360	9 AC021105	AL088243 Human DNA
C 14	439.6	4.4	176287	9 HSBAL2M9	AL088243 Human DNA
C 15	439.4	4.4	8852	10 WEMPEX1	X36829 M.musculu
C 16	436.2	4.4	185909	9 AL590434	AL590434 Human DNA
C 17	430.2	4.3	248281	9 AC008737	AC008737 Homo sapi
C 18	429	4.3	275982	2 AC131626	AL131626 Rattus no
C 19	428	4.3	175827	2 AL358472	AL358472 Human DNA
C 20	422.4	4.2	182535	9 AC079331	AC079331 Homo sapi
C 21	422.4	4.2	216789	9 AC006534	AC006534 Homo sapi
C 22	419.4	4.2	457	9 HDMXMA02	M36798 Human hemop
C 23	415.8	4.2	148525	2 AC016675	AC016675 Homo sapi
C 24	414.6	4.1	178105	9 AC073057	AC073057 Homo sapi
C 25	414.2	4.1	57722	2 HSCB33B10	282175 Homo sapien
C 26	410.2	4.1	106008	9 HSE06	AL020993 Human DNA
C 27	408.6	4.1	160822	2 AC020920	AC020920 Homo sapi
C 28	408.4	4.1	193148	9 AP003732	AP003732 Homo sapi
C 29	405.8	4.1	36095	9 AC013256	AC013256 Homo sapi
C 30	404.4	4.0	190570	2 AL157894	AL157894 Homo sapi
C 31	403.6	4.0	185559	2 AC019166	AC019166 Homo sapi
C 32	403.4	4.0	102717	9 AC004079	AC004079 Homo sapi
C 33	403.4	4.0	164026	2 AC092762	AC092762 Pan trogl
C 34	403	4.0	286758	9 AC006449	AC006449 Homo sapi
C 35	401.6	4.0	187383	2 AC138954	AC138954 Homo sapi
C 36	400.6	4.0	73431	2 AP000575	AP000575 Homo sapi
C 37	400.4	4.0	155824	2 AC087296	AC087296 Homo sapi
C 38	399.6	4.0	191440	9 AC012645	AC012645 Homo sapi
C 39	399	4.0	61993	9 AP000828	AP000828 Homo sapi
C 40	398.2	4.0	146741	9 AC011452	AC011452 Homo sapi
C 41	398	4.0	195880	9 AC074117	AC074117 Homo sapi
C 42	397.4	4.0	178459	2 AC079931	AC079931 Homo sapi
C 43	397	4.0	116478	9 AC092287	AC092287 Homo sapi
C 44	397	4.0	172876	9 AC114489	AC114489 Homo sapi
C 45	397	4.0	179221	9 AC114493	AC114493 Homo sapi

ALIGNMENTS

RESULT 1
AC068733 191656 bp DNA linear PRI 24-MAY-2002
LOCUS Homo sapiens chromosome 11, clone RP11-304C12, complete sequence.
DEFINITION AC068733
ACCESSION AC068733
VERSION AC068733.12 GI:19703148
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 191656)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,B.
TITLE Homo sapiens chromosome 11, clone RP11-304C12
JOURNAL Unpublished

Pred. No. is the number of results predicted by chance to have a

THIS PAGE BLANK (uspto)

;; CURRENT FILING DATE: 2002-04-26
;; PRIOR APPLICATION NUMBER: US 60/286,764
;; PRIOR FILING DATE: 2001-04-26
;; NUMBER OF SEQ ID NOS: 80
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 5
;; LENGTH: 606
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-133-797-5

Query Match 9.1%; Score 199; DB 15; Length 606;
Beet Local Similarity 25.0%; Pred. No. 3.2e-10;
Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

```
QY 5 LGAPVALGL-----WSLCMSLAIAATPL-----PPTSAHGUNAEGETKPDPOVTERCS 51
Db 277 VGDFLRVGLPYEDKRVWQLYGVRESVSPITQPEEPFLPEPPDNRSSAPPRKDVPHRCS 336
QY 52 DGWSPDATTTLDNGTMLFFKGEFVWKSHKMDRELISER-----WKNFP--SPVDAAF 101
Db 337 T--HPDAVA-QIRGEAFFKGYFWRLLTR-DRLVSLQPAQMRHFRGLPLHDSVDAY 392
QY 102 -RQHNVSFLIKGDKWVYTPPEKKEGYPKLLQDEFPGIPSPLDAAVECHRGECQAEGL 160
Db 393 ERTSDHKIVFFKGDRIYWFKDNVNEBEGYRPVSD-FSLPPGIDAA----- 437
QY 161 FFOGHGHRNGTGHGNSSTHGPYWRCSPHLVLSALTSDNHGATYAFSGTHYRLDT--SRD 219
Db 438 FSNMHNDR-----TYFKDQLYWRYYDHTRH 463
QY 220 GWHSMPI-AHQWPOGSPSAVDAAFSNEB-KLYLVQSTQVYVFLTKGGYTLVSGYPRLEKE 277
Db 464 MDPGYPAQSPIMRGVPSITLDDAMRMMSDGSYFFRG-QEYWKVLDGELEVAAPGYPOSTARD 522
QY 278 -----VGTPHGIIIDSVDAA 292
Db 523 WLVCGDSQADGSVAAGVDAA 542
```

Search completed: December 16, 2003, 06:59:58
Job time : 253 secs

Db 250 T--HFDVA-QIRGEAFPFKGYFWRLTR-DRHLVSLQPAQMRFWRGRLPLHDSVDVAY 305
QY 102 -RGHNSVFLIKGDKVWVYPRPEKKEGYPKLLODEFFGIPSPIDAVECHRGECQAEVYL 160
Db 306 ERTSDHKIYFFKGDYRWVFEDNNVEEGYPRPVSDFSLPPGIGIDA----- 350
QY 161 FFOGHGRNCTGCHNSTHGHPEYMRCSPHLVLSALTSNDHGATYAFSGTHYWRDLT-SRD 219
Db 351 FSWAHNDR-----TYFFKQDLVWRYDDHTRH 376
QY 220 GWHSMPI-AHQWPGPSAVDAAFSWEK-KLYLVQGTQVYVFLTKGGYTLVSGYPKRLEKE 277
Db 377 MDGYPAGSPLMKGVSTLDDARMSDGSAYFFRG-QEYWKVLDGELEVAAPGYPOSTARD 435
QY 278 -----VGTPIGIIILDSVDA 292
Db 436 WLVCGDSQADGSVAAGVDA 455

RESULT 13

US-10-131-985-49
; Sequence 49, Application US/10131985
; Publication No. US2003019440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Ocleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 49
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-49

Query Match 9.1%; Score 199; DB 12; Length 519;
Best Local Similarity 25.0%; Pred. No. 2.6e-10;

Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

QY 5 LGAPVALGL-----WSLCWSLAIATPL-----PPTSAHGNAVBGETKPPDDVTERCS 51
Db 190 VGOPLRYGLYEDKVRWQLYGVRESVSPYAPQPEEPPLPEPPDNRSSAPPRKDVPHRCS 249
QY 52 DGWSPDATTLDNGTMLPFKGEFVWMSKMDRELISER-----WKNFP---SPVDAAF 101
Db 250 T--HFDVA-QIRGEAFPFKGYFWRLTR-DRHLVSLQPAQMRFWRGRLPLHDSVDVAY 305
QY 102 -RGHNSVFLIKGDKVWVYPRPEKKEGYPKLLODEFFGIPSPIDAVECHRGECQAEVYL 160
Db 306 ERTSDHKIYFFKGDYRWVFEDNNVEEGYPRPVSDFSLPPGIGIDA----- 350
QY 161 FFOGHGRNCTGCHNSTHGHPEYMRCSPHLVLSALTSNDHGATYAFSGTHYWRDLT-SRD 219
Db 351 FSWAHNDR-----TYFFKQDLVWRYDDHTRH 376
QY 220 GWHSMPI-AHQWPGPSAVDAAFSWEK-KLYLVQGTQVYVFLTKGGYTLVSGYPKRLEKE 277
Db 377 MDGYPAGSPLMKGVSTLDDARMSDGSAYFFRG-QEYWKVLDGELEVAAPGYPOSTARD 435
QY 278 -----VGTPIGIIILDSVDA 292
Db 436 WLVCGDSQADGSVAAGVDA 455

RESULT 14

US-10-406-209-2
; Sequence 2, Application US/10406209
; Publication No. US20030170758A1
; GENERAL INFORMATION:
; APPLICANT: KYOMA HAKKO KOGYO CO., LTD.
; TITLE OF INVENTION: NOVEL ANTIBODIES, DRUGS CONTAINING THESE ANTIBODIES AND
; TITLE OF INVENTION: METHODS FOR
; FILE REFERENCE: 1241.19
; CURRENT APPLICATION NUMBER: US/10/406,209
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/806,228C
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/JP99/05350
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 10-291501
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 10-291503
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-406-209-2

Query Match 9.1%; Score 199; DB 12; Length 605;
Best Local Similarity 25.0%; Pred. No. 3.2e-10;

Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

QY 5 LGAPVALGL-----WSLCWSLAIATPL-----PPTSAHGNAVBGETKPPDDVTERCS 51
Db 276 VGOPLRYGLYEDKVRWQLYGVRESVSPYAPQPEEPPLPEPPDNRSSAPPRKDVPHRCS 335
QY 52 DGWSPDATTLDNGTMLPFKGEFVWMSKMDRELISER-----WKNFP---SPVDAAF 101
Db 336 T--HFDVA-QIRGEAFPFKGYFWRLTR-DRHLVSLQPAQMRFWRGRLPLHDSVDVAY 391
QY 102 -RGHNSVFLIKGDKVWVYPRPEKKEGYPKLLODEFFGIPSPIDAVECHRGECQAEVYL 160
Db 392 ERTSDHKIYFFKGDYRWVFEDNNVEEGYPRPVSDFSLPPGIGIDA----- 436
QY 161 FFOGHGRNCTGCHNSTHGHPEYMRCSPHLVLSALTSNDHGATYAFSGTHYWRDLT-SRD 219
Db 437 FSWAHNDR-----TYFFKQDLVWRYDDHTRH 462
QY 220 GWHSMPI-AHQWPGPSAVDAAFSWEK-KLYLVQGTQVYVFLTKGGYTLVSGYPKRLEKE 277
Db 463 MDGYPAGSPLMKGVSTLDDARMSDGSAYFFRG-QEYWKVLDGELEVAAPGYPOSTARD 521
QY 278 -----VGTPIGIIILDSVDA 292
Db 522 WLVCGDSQADGSVAAGVDA 541

RESULT 15

US-10-133-797-5
; Sequence 5, Application US/10133797
; Publication No. US20030109021A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Shujian
; APPLICANT: Chen, Jian
; APPLICANT: Feder, Liana
; APPLICANT: Lee, Liana
; APPLICANT: Krystek, Stanley
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY
; FILE REFERENCE: D0141NP
; CURRENT APPLICATION NUMBER: US/10/133,797

QY 202 ATYASGTHYRLDTSRD---GHSWPIAHQWPGPSAVDAFSWEK-LYLVQGTQVY 256
Db 392 -TYFVDNQWRYDERBROQMDPGYFKLITKNFGGIGPK-IDAVFYSKNXYFFQGSNOF 449
QY 257 VF 258
Db 450 EY 451

RESULT 10

US-10-171-311-135
; Sequence 135, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatz, Karen
; APPLICANT: Ganavaram, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-135

Query Match 9.3%; Score 202.5; DB 15; Length 470;
Best Local Similarity 27.3%; Pred. No. 1.1e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAGETKPPDVTY--RCSGWSFDATLDDNGTMLFKGEFVW---KSHKWR 83
Db 260 SLYGDPKRNQRLPNPNDSEPALCDPNLSFDVAVTVGN-KIFFKDRFFWLKVSERPKTSV 318
QY 84 ELISERWKNPSPVDAFR--OGHNSVFLIKGDKVWYPPPEKKEKGYPKLQD-EPPGIPS 141
Db 319 NLISLSLWPTLPSGIEAIEIARNOVFLPKODKYWLISNLAPENYPSISFGFBNVVK 378
QY 142 PLDAVBECHRECAEGVLFFQGHGRNGTGHGNSTHGPEYMCSPHLVLSALTSQNHG 201
Db 379 KIDAAV-----FNPRFYR----- 391
QY 202 ATYAFSGTHYRLDTSRD---GHSWPIAHQWPGPSAVDAFSWEK-LYLVQGTQVY 256
Db 392 -TYFVDNQWRYDERBROQMDPGYFKLITKNFGGIGPK-IDAVFYSKNXYFFQGSNOF 449
QY 257 VF 258
Db 450 EY 451

RESULT 11
US-09-925-301-1217
; Sequence 1217, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1217
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1217

Query Match 9.3%; Score 202.5; DB 9; Length 473;
Best Local Similarity 27.3%; Pred. No. 1.1e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAGETKPPDVTY--RCSGWSFDATLDDNGTMLFKGEFVW---KSHKWR 83
Db 263 SLYGDPKRNQRLPNPNDSEPALCDPNLSFDVAVTVGN-KIFFKDRFFWLKVSERPKTSV 321
QY 84 ELISERWKNPSPVDAFR--OGHNSVFLIKGDKVWYPPPEKKEKGYPKLQD-EPPGIPS 141
Db 322 NLISLSLWPTLPSGIEAIEIARNOVFLPKODKYWLISNLAPENYPSISFGFBNVVK 381
QY 142 PLDAVBECHRECAEGVLFFQGHGRNGTGHGNSTHGPEYMCSPHLVLSALTSQNHG 201
Db 382 KIDAAV-----FNPRFYR----- 394
QY 202 ATYAFSGTHYRLDTSRD---GHSWPIAHQWPGPSAVDAFSWEK-LYLVQGTQVY 256
Db 395 -TYFVDNQWRYDERBROQMDPGYFKLITKNFGGIGPK-IDAVFYSKNXYFFQGSNOF 452
QY 257 VF 258
Db 453 EY 454

RESULT 12
US-09-801-196-30
; Sequence 30, Application US/09801196
; Patent No. US20020037827A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Kai
; APPLICANT: Smith, Ryan
; APPLICANT: Fajardo, Mark
; APPLICANT: Moss, Patrick
; TITLE OF INVENTION: A NOVEL MATRIX METALLOPROTEINASE (MMP-25)
; TITLE OF INVENTION: EXPRESSED IN SKIN CELLS
; FILE REFERENCE: 240083.509
; CURRENT APPLICATION NUMBER: US/09/801,196
; CURRENT FILING DATE: 2001-03-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-801-196-30

Query Match 9.1%; Score 199; DB 9; Length 519;
Best Local Similarity 25.0%; Pred. No. 2.6e-10;
Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

QY 5 LGAPVALGL-----MSLCWSLAIATPL-----PPTSAGNVAGETKPPDPTVTERCS 51
Db 190 VGDPILRYGLPREDKVRWOLYGVRESVSPTAQPEFPLLEBPDPNBSAPPRKDVPRCS 249
QY 52 DGWSFDATLDDNGTMLFKGEFVWKSMDRELISER-----WKNFP---SPVDAF 101

```

; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-337-632-2

Query Match          9.3%; Score 202.5; DB 12; Length 470;
Beet Local Similarity 27.3%; Pred. No. 1.1e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10.

QY   30 SAHGNVAEGETKPPDPVTE--RCSGSGSPDATLLDNGTMLFFKGEEFW----KSHKKDR 83
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   260 SLVDPPENQRPLRPDNSEPALCDPNSFDALTTCN-KIFFKDRFWMKVSRPKTSV 318

QY   84 ELISERNKNFPSPVDAAF-R-QGHNSVFLLIKDKWVYPPEKEKGYPRLQD-EPPGI PS 141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   319 NLISLPTLPISGIEAAYEIRARQVFLFKDDKWMLSLNLAPRPYPSGISHSFGPNVK 378

QY   142 PLDAAVECHRGCEAGVLFFQGHGRKGTSGNSTHHGPFTMCRSPALVLSALTSDNHG 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   379 KIDAAV-----FNPRFYR----- 391

QY   202 ATYAFCSTHYWRDLTSDR---GMHSNPILAHQWPQSADVAAAFSWEKK-LYLVOGTQVY 256
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   392 -TYFFVDNQWRYDERQMDPGYPKLTITKNGSIGPK-IDAVFSKMKYYTFFQGSNOF 449

QY   257 VF 258
      :
Db   450 EF 451

```

```

US-10-131-985-39
: Sequence 39. Application US/10131985
: Publication No. US20030199440A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Dack, Kevin N
: APPLICANT: Davies, Michael J
: APPLICANT: Fish, Paul V
: APPLICANT: Huggins, Jonathan P
: APPLICANT: McIntosh, Fraser S
: APPLICANT: Occleston, Nicholas L
: TITLE OF INVENTION: Composition
:
: FILE REFERENCE: PCS 10391A
:
: CURRENT APPLICATION NUMBER: US/10/131,985
:
: CURRENT FILING DATE: 2002-04-25
:
: PRIOR APPLICATION NUMBER: US/09/726,295
:
: PRIOR FILING DATE: 2000-11-30
:
: PRIOR APPLICATION NUMBER: GB 9930768.8
:
: PRIOR FILING DATE: 1999-12-29
:
: NUMBER OF SEQ ID NOS: 60
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 39
:
: LENGTH: 470
:
: TYPE: PRT

```

Query Match	9.3%	Score 202.5	DB 12	Length 470
Best Local Similarity	27.3%	Pred. No. 1,1e-10		
Matches	66	Conservative 23	Mismatches 90	Indels 63
				Gaps 10
Qy	30	SAHGNVAGEGKPPDPVTE--RCSDGSPDATTLDNGTLMFPKGEFVW-----KSHKMDR	83	
		: : : : : : : : : : : : : : : : : : :		
Db	260	SLVGDPENENORLPDPDSEPALDCPNISFDPAVTTVGR-KLFFKDRPFMYLKVSRPKTSV	318	
		: : : : : : : : : : : : : : : : : : :		
Qy	84	ELISERKMNPPSPYDAFR-QGHNSEVLKIGDKKVVYPPEKEKGYPPLD-EPFGIPS	141	
		: : : : : : : : : : : : : : : : : : :		
Db	319	NLISLSTPPLPSGIEAAYEIEARQVPLFKDDKWLISNLRPBENYPSKSIHSPGPNPVK	378	
		: : : : : : : : : : : : : : : : : : :		
Qy	142	PLDAAVECHRGEOCAEGLVFPQGHGHNKGTHGNGSTHGGREYMGCSPLVLVALTSDNHG	201	
		: : : : : : : : : : : : : : : : : : :		
Db	379	KIDAAV-----FNPRFYR-----	391	

PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/JP99/05350
PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 10-291501
PRIOR FILING DATE: EARLIER FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 10-291503
PRIOR FILING DATE: EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 28
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1
LENGTH: 587
TYPE: PRT
ORGANISM: Mouse
US-10-406-209-1

Query Match 9.3%; Score 204; DB 12; Length 587;
Best Local Similarity 24.4%; Pred. No. 1e-10; Indels 158; Gaps 25;
Matches 103; Conservative 34; Mismatches 127;

QY 5 LGARVALGL-----WSLCMSLAATPL-----PPTSANGVAGETKPDV 46
DB 273 VGDVRYGLPYEDRYVWQLYGVRESVPTAQLOTPEPEEPPLPEPPNNSSTPPQKV 332
QY 47 TERCSGKSPDATTLDNGTMLPFKGEFVWKSHPKDRLEISER-----WKNP--SP 96
DB 333 PHRCTA--HFDVAVA-QIRGEAFFEKGYFWRLTR--DRHLVSLQPAQMERFWRGJPLHDS 388
QY 97 VDAAF-RGSHSVFLIKGDKVWVYVPEKKEGYPLLODEFPGLSPILDAVECHRGSCQ 155
DB 389 VDAYERSTSHKIVFPGKDRYVWFQDNNEEGYPRVSD--PSLPGGIDA-----437
QY 156 AEGVLFQGHGRNGTGHGNGSTHGRPEYMRCSPHLVLSALTSDNGATVAFSGTHYMRLD 215
DB 438 ----VFSANHDR-----TYFFKQDQLYMYXD 459
QY 216 --TSR--DGHMS-MPIAHQWPGPSAADAAPSWEK-KLYLVQGTQVYVFLTKGTYTVSG 269
DB 460 DHTRMDFGYPAQGL--WRGVSMMLDAMRWSDGASYFRFG--QEWKVLDDGELEAAG 515
QY 270 YPKLEKE--VGTPHGIIIDSVDAAFICPGSSRHIMAGRIMWLDLKSQAQATWETLP 326
DB 516 YPQSTARDWLVCGE-----LADAEVGPBPQ-----GR-----SQAQ-----548
QY 327 WPHEKVDG-ALCMEXSLGPNCSANGPGLYLHGNLYCYSDVEKLANAKALPOQVNTS 385
DB 549 -----DGLAVCS-----CTSDARL-----ALPSLLLTLP 573
QY 386 LL 387
DB 574 LL 575

RESULT 5
US-09-801-196-23
Sequence 23, Application US/09801196
Patent No. US20020037827A1
GENERAL INFORMATION:
APPLICANT: Wang, Kai
APPLICANT: Smith, Ryan
APPLICANT: Fajardo, Mark
APPLICANT: Moss, Patrick
TITLE OF INVENTION: A NOVEL MATRIX METALLOPROTEINASE (MMP-25)
FILE REFERENCE: 240083.509
CURRENT APPLICATION NUMBER: US/09/801.196
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 470
TYPE: PRT
ORGANISM: Homo sapiens
US-09-801-196-23

Query Match 9.3%; Score 202.5; DB 9; Length 470;
Best Local Similarity 27.3%; Pred. No. 1.1e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAGETKPPDYTE--RCSGWSFDAATLDDNGTMLPFKGEFVW-----KSHKMDR 83
DB 260 SLVGDPEKNQRLPNDSEBPCDNLSPDAVTVGN-KIFFKDRFFWLKVSRPKTSV 318
QY 84 ELISERWKNFSPYDAAFR-QGHSVFLIKGDKVWVYVPEKKEGYPLLOD-EFPGIPS 141
DB 319 NLISLWPTLPSCGLEAIVEIARNQVFLPKDKTWLISNLPBEPNYPKSHSFGFPNVK 378
QY 142 PLDAVECHRGCEQAEGLVLFQGHGRNGTGHGNGSTHGRPEYMRCSPHLVLSALTSDNHG 201
DB 379 KIDAAV-----FNPRFYR-----391
QY 202 ATYAFSGTHWRDLTSD--GHSWPIAHQWPGPSAADAAPSWEK-LYLVQGTQVY 256
DB 392 -TYFFVNDQYWRDYERQMDPGYPKLITKFGIGPK-IDAVFYSKKYYVFFQSGNQF 449
QY 257 VF 258
DB 450 EY 451

RESULT 6
US-09-920-455-219
Sequence 219, Application US/09920455
Patent No. US20020168647A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.540
CURRENT APPLICATION NUMBER: US/09/920.455
CURRENT FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 275
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 219
LENGTH: 470
TYPE: PRT
ORGANISM: Homo sapiens
US-09-920-455-219

Query Match 9.3%; Score 202.5; DB 10; Length 470;
Best Local Similarity 27.3%; Pred. No. 1.1e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAGETKPPDYTE--RCSGWSFDAATLDDNGTMLPFKGEFVW-----KSHKMDR 83
DB 260 SLVGDPEKNQRLPNDSEBPCDNLSPDAVTVGN-KIFFKDRFFWLKVSRPKTSV 318
QY 84 ELISERWKNFSPYDAAFR-QGHSVFLIKGDKVWVYVPEKKEGYPLLOD-EFPGIPS 141
DB 319 NLISLWPTLPSCGLEAIVEIARNQVFLPKDKTWLISNLPBEPNYPKSHSFGFPNVK 378
QY 142 PLDAVECHRGCEQAEGLVLFQGHGRNGTGHGNGSTHGRPEYMRCSPHLVLSALTSDNHG 201
DB 379 KIDAAV-----FNPRFYR-----391
QY 202 ATYAFSGTHWRDLTSD--GHSWPIAHQWPGPSAADAAPSWEK-LYLVQGTQVY 256
DB 392 -TYFFVNDQYWRDYERQMDPGYPKLITKFGIGPK-IDAVFYSKKYYVFFQSGNQF 449
QY 257 VF 258
DB 450 EY 451

RESULT 7
US-10-301-822-123

```

Db      181 PEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPIAHQWPGPSAVDA 240
Qy      241 FSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGIIIDSVDAATICGSSR 300
Db      241 FSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGIIIDSVDAATICGSSR 300
Qy      301 LHTMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNSCSANGPGIYLHGP 360
Db      301 LHTMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNSCSANGPGIYLHGP 360
Qy      361 NLICYSDVEKLNAAKALPOPONVTSILGCTH 391
Db      361 NLICYSDVEKLNAAKALPOPONVTSILGCTH 391

```

```

RESULT 2
US-09-900-448-4
; Sequence 4, Application US/09900448
; Publication No. US20030220488A1
; GENERAL INFORMATION:
; APPLICANT: CECCARDI, Toni et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001272
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Human
US-09-900-448-4

```

```

Query Match      97.9%; Score 2139.5; DB 12; Length 462;
Best Local Similarity 84.6%; Pred. No. 1.2e-196;
Matches 391; Conservative 0; Mismatches 0; Indels 71; Gaps 1;

```

```

Qy      1 1MAVLCAPVALGLMSLCSLAIAITPLPPTSAGHNVAGE--TKPDDVTERGSDGMSFPA 60
Db      1 1MAVLCAPVALGLMSLCSLAIAITPLPPTSAGHNVAGE--TKPDDVTERGSDGMSFPA 60
Qy      61 LDDNGTMLPFKGFVWMSHKMDRELISERKKNFSPVDAAPROGNSVFLIKGDKWVYP 120
Db      61 LDDNGTMLPFKGFVWMSHKMDRELISERKKNFSPVDAAPROGNSVFLIKGDKWVYP 120
Qy      121 PEKKEKGYPKLLQDEFPGISPLDAVECHRGECQAEGLVFFQ----- 163
Db      121 PEKKEKGYPKLLQDEFPGISPLDAVECHRGECQAEGLVFFQ----- 163
Qy      164 -----GHCGRN 169
Db      164 -----GHCGRN 169
Qy      181 SWPVGNCSSALRWLGRRYCFQGNQFLRFPVGEVPRRYRVDYFEMPCGRGHGRN 240
Db      181 SWPVGNCSSALRWLGRRYCFQGNQFLRFPVGEVPRRYRVDYFEMPCGRGHGRN 240
Qy      170 GTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPIAHQ 229
Db      170 GTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPIAHQ 229
Qy      241 GTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPIAHQ 300
Db      241 GTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPIAHQ 300
Qy      230 WPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGIIIDSV 289
Db      230 WPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGIIIDSV 289
Qy      301 WPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGIIIDSV 360
Db      301 WPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGIIIDSV 360
Qy      290 DAAPICGSSRLHIMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNSCSA 349
Db      290 DAAPICGSSRLHIMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNSCSA 349
Qy      361 DAAPICGSSRLHIMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNSCSA 420
Db      361 DAAPICGSSRLHIMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNSCSA 420
Qy      350 NGPGLYLIHGPNLYCYSDEVEKLNAAKALPOPONVTSILGCTH 391
Db      421 NGPGLYLIHGPNLYCYSDEVEKLNAAKALPOPONVTSILGCTH 462

```

RESULT 3

```

US-10-316-253-40
; Sequence 40, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Peng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-40

```

```

Query Match      71.2%; Score 1556.5; DB 12; Length 460;
Best Local Similarity 63.8%; Pred. No. 1.1e-140;
Matches 296; Conservative 35; Mismatches 54; Indels 79; Gaps 6;

```

```

Qy      1 1MAVLCAPVALGLMSLCSLAIAITPLPPTSAGHNVAGE--TKPDDVTERGSDGMSFPA 58
Db      1 1MAVLCAPVALGLMSLCSLAIAITPLPPTSAGHNVAGE--TKPDDVTERGSDGMSFPA 58
Qy      59 TTLDDNGTMLPFKGFVWMSHKMDRELISERKKNFSPVDAAPROGNSVFLIKGDKWV 118
Db      59 TTLDDNGTMLPFKGFVWMSHKMDRELISERKKNFSPVDAAPROGNSVFLIKGDKWV 118
Qy      119 YPEKKEKGYPKLLQDEFPGISPLDAVECHRGECQAEGLVFFQCH----- 165
Db      119 YPEKKEKGYPKLLQDEFPGISPLDAVECHRGECQAEGLVFFQCH----- 165
Qy      166 -----GH-- 167
Db      166 -----GH-- 167
Qy      178 ERSWPAVNGCTALRWLGRRYCFQGNKFLRFPVGEVPRRYRVDYFISCPGRGHGK 237
Db      178 ERSWPAVNGCTALRWLGRRYCFQGNKFLRFPVGEVPRRYRVDYFISCPGRGHGK 237
Qy      168 -RNGTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPI 226
Db      168 -RNGTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPI 226
Qy      238 LKNGTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPI 295
Db      238 LKNGTGHGNSHTHGPPEYMRCSPLVLVSALTSDNHGATYAFTSGTHYMRDLTSDRGMSWPI 295
Qy      227 AHQWPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGII 286
Db      227 AHQWPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGII 286
Qy      296 AHQWPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGII 355
Db      296 AHQWPGPSAVDAAFSWEEELLYVQGVVFLTKGTYLVSGYPRKLEKEVGTGPHGII 355
Qy      287 DSVDAAPICGSSRLHIMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNS 346
Db      287 DSVDAAPICGSSRLHIMAGRRLMWDLKSGAQTWTELPMPHEKVDGALCMESLGPNS 346
Qy      347 CSANGPGLYLIHGPNLYCYSDEVEKLNAAKALPOPONVTSILGCTH 390
Db      347 CSANGPGLYLIHGPNLYCYSDEVEKLNAAKALPOPONVTSILGCTH 390
Qy      416 CSANGPGLYLIHGPNLYCYSDEVEKLNAAKALPOPONVTSILGCTH 459
Db      416 CSANGPGLYLIHGPNLYCYSDEVEKLNAAKALPOPONVTSILGCTH 459

```

```

RESULT 4
US-10-406-209-1
; Sequence 1, Application US/10406209
; Publication No. US20030170758A1
; GENERAL INFORMATION:
; APPLICANT: KYOMA HAKKO KOGYO CO., LTD.
; TITLE OF INVENTION: NOVEL ANTIBODIES, DRUGS CONTAINING THESE ANTIBODIES AND
; TITLE OF INVENTION: METHODS FOR
; FILE REFERENCE: 1241.19
; CURRENT APPLICATION NUMBER: US/10/406,209
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/806,228C

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:48:58 ; Search time 252 Seconds
(without alignments)
288.570 Million cell updates/sec

Title: US-09-900-448-2
Perfect score: 2185
Sequence: 1 MARVLGAPVALGLMSLCWSL.....NAKALPQPNVTSLLGCTH 391

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues
Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2185	100.0	391	12	US-09-900-448-2
2	2139.5	97.9	462	12	US-09-900-448-4
3	1556.5	71.2	460	12	US-10-316-253-40
4	204	9.3	587	12	US-10-406-209-1
5	202.5	9.3	470	9	US-09-801-196-23
6	202.5	9.3	470	10	US-09-920-455-219
7	202.5	9.3	470	12	US-10-301-822-123
8	202.5	9.3	470	12	US-10-337-632-2
9	202.5	9.3	470	12	US-10-131-985-39
10	202.5	9.3	470	15	US-10-171-311-135
11	202.5	9.3	473	9	US-09-925-301-1217
12	199	9.1	519	9	US-09-801-196-30
13	199	9.1	519	12	US-10-131-985-49
14	199	9.1	605	12	US-10-406-209-2
15	199	9.1	606	15	US-10-133-797-5

16	195.5	8.9	477	10	US-09-920-455-260	Sequence 260, App
17	189.5	8.7	476	9	US-09-801-196-25	Sequence 25, App1
18	189.5	8.7	476	12	US-10-021-660-72	Sequence 72, App1
19	189.5	8.7	476	12	US-10-131-985-35	Sequence 35, App1
20	183.5	8.4	467	9	US-09-801-196-25	Sequence 20, App1
21	183.5	8.4	467	11	US-09-759-1308-176	Sequence 176, App
22	183.5	8.4	467	12	US-10-131-985-31	Sequence 31, App1
23	181.5	8.3	469	9	US-09-801-196-19	Sequence 19, App1
24	181.5	8.3	469	9	US-09-853-386-100	Sequence 100, App
25	181.5	8.3	469	12	US-10-301-822-119	Sequence 119, App
26	181.5	8.3	469	12	US-10-021-660-76	Sequence 76, App1
27	181.5	8.3	469	12	US-10-308-279-34	Sequence 34, App1
28	181.5	8.3	469	12	US-10-131-985-23	Sequence 23, App1
29	181.5	8.3	469	15	US-10-106-698-6283	Sequence 6283, Ap
30	179.5	8.2	488	9	US-09-801-196-26	Sequence 26, App1
31	179.5	8.2	488	12	US-10-301-822-121	Sequence 121, App
32	179.5	8.2	488	12	US-10-131-985-37	Sequence 37, App1
33	179.5	8.2	488	15	US-10-177-293-298	Sequence 298, App
34	179.5	8.2	488	16	US-10-225-486-67	Sequence 67, App1
35	177	8.1	477	9	US-09-801-196-24	Sequence 24, App1
36	177	8.1	477	12	US-10-301-822-127	Sequence 127, App
37	177	8.1	477	12	US-10-131-985-27	Sequence 27, App1
38	177	8.1	477	15	US-10-171-311-137	Sequence 137, App
39	176	8.1	471	9	US-09-801-196-31	Sequence 31, App1
40	176	8.1	471	9	US-09-801-196-32	Sequence 32, App1
41	176	8.1	471	12	US-10-375-884-1	Sequence 1, App1
42	176	8.1	471	12	US-10-375-884-21	Sequence 21, App1
43	176	8.1	471	12	US-10-131-985-41	Sequence 41, App1
44	176	8.1	471	15	US-10-050-2168-6	Sequence 6, App1
45	176	8.1	471	15	US-10-075-069-1	Sequence 1, App1

ALIGNMENTS

RESULT 1
US-09-900-448-2
; Sequence 2, Application US/09900448
; Publication No. US20030220488A1
GENERAL INFORMATION:
; APPLICANT: CECARDI, Toni et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001272
; CURRENT APPLICATION NUMBER: US/09/900,448
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Human
US-09-900-448-2

Query Match	100.0%;	Score 2185;	DB 12;	Length 391;
Best Local Similarity	100.0%;	Pred. No. 4e-201;		
Matches 391;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MARVLGAPVALGLMSLCWSLIAITPLPTSAHGNVAGETKPPDPVTERCSGMSFDATT	60	
Db	1	MARVLGAPVALGLMSLCWSLIAITPLPTSAHGNVAGETKPPDPVTERCSGMSFDATT	60	
Qy	61	LDNGTMTLFFKSGFVWKSMDRELISERKKNPPSPDAAFRGHNSVFLIKDKVWVP	120	
Db	61	LDNGTMTLFFKSGFVWKSMDRELISERKKNPPSPDAAFRGHNSVFLIKDKVWVP	120	
Qy	121	PEKKEKYPKLLDDEFFGISPLDAAVECHRGSCQAGVLFPOGHGRNGTGHNSTHHG	180	
Db	121	PEKKEKYPKLLDDEFFGISPLDAAVECHRGSCQAGVLFPOGHGRNGTGHNSTHHG	180	
Qy	181	PEYWRCSPHLVLSALTSNDHGAATYAFSGTHYWRLDTSRDSGMSWPIAHQWPGSPSAVDAA	240	
Db	181	PEYWRCSPHLVLSALTSNDHGAATYAFSGTHYWRLDTSRDSGMSWPIAHQWPGSPSAVDAA	240	

THIS PAGE BLANK (USPTO)

```

Db      266 TPSTPTLSSQAGTDTNEIALQEEVPEVCET--SFDAVS--TIRGELFFFKAGFVRL 322
Qy      79 HKMDRE-----LISERWKNFPPSPVDAAFROGHNSVFLIKGDKVWVYPEPEKKEK 133
Db      323 RSGQLQPGVYPALASHRMQGLPSVDAAFEDAQGIWFGQAQYVYDEKEFLGPAFLSK 382
Qy      134 DEFPGIPSPDLDAVECHRGCEQAGVLFPGHGHNRNGTGHNGSTHGPYRCSPHLVLS 193
Db      383 LGLGQ--SPVHAALV-----WGPE----- 399
Qy      194 ALTSNHNATYAFSGTHYWRDL--TSR--DGMHSPVIAHQ--WPGPSAVDAAFSWEKL 247
Db      400 -----KKKIYFFRGDYMRFHPTQRYVDN--FVPRRTDWRGVPSEIDAAF----- 443
Qy      248 YLVQGTQVYVFLTKG 262
Db      444 ---QDADGVAYFLRG 452

```

RESULT 15

```

S13423
N:Altermate names: EC 3.4.24.-) Precursor - human
C:Species: Homo sapiens (man)
C:Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 22-Jun-1999
C:Accession: S13423; 138250; S58912
R:Basset, P.; Belllocq, J.P.; Wolf, C.; Stoll, I.; Hutin, P.; Limacher, J.M.; Podhajcer,
Nature 348, 699-704, 1990
A:Title: A novel metalloproteinase gene specifically expressed in stromal cells of breast
A:Reference number: S13423; MUID:91080920; PMID:1701851
A:Accession: S13423
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-488 <BAS>
A:Cross-references: EMBL:X57766
R:Anglard, P.; Melot, T.; Guerin, E.; Thomas, G.; Basset, P.
J. Biol. Chem. 270, 20337-20344, 1995
A:Title: Structure and promoter characterization of the human stromelysin-3 gene.
A:Reference number: 138250; MUID:95386471; PMID:7657606
A:Accession: 138250
A:Status: Preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-36 <RES>
A:Cross-references: EMBL:X84664; NID:G984746; PIDN:CAA59150.1; PID:G987949
R:Pei, D.; Weiss, S.J.
Nature 375, 244-247, 1995
A:Title: Furin-dependent intracellular activation of the human stromelysin-3 zymogen.
A:Reference number: S58912; MUID:95265105; PMID:7746327
A:Accession: S58912
A:Status: Preliminary
A:Molecule type: Protein
A:Residues: 81-101 <PEI>
C:Genetics:
A:Gene: GDB:MMP11; STWY3
A:Cross-references: GDB:128630; OMIM:185261
A:Map position: 22q11.2-22q11.2
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotei
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:48-256/Domain: matrix metalloproteinase homology <MMP>
F:80-215,219,225/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status
F:215,219,225/Binding site: zinc, catalytic (His) (active) #status predicted
F:216/Active site: Glu #status predicted

```

Query Match

Best local similarity 8.2%; Score 179.5; DB 2; Length 488;
Matches 65; Conservative 23; Mismatches 100; Indels 71; Gaps 12;

```

Qy      13 LMSLCWGLATITPPLPSAHGNVAEGTKP--DPVTERCSGWSFDPATITLDNGTMLPFK 71
Db      256 LVGGPWP-TVTSRTPALGPAGIDTNETIAPLEBPADPACBA-SFDAVS--TIRGELFFFK 312

```

```

Qy      72 GEPYWKSHKMDRE-----LISERWKNFPPSPVDAAFROGHNSVFLIKGDKVWVYPEPEKKEK 126
Db      313 AGFVWRLRGQGLQGYPALASHRMQGLPSVDAAFEDAQGIWFGQAQYVYDEKEFL 372
Qy      127 GYPLLDDEFPGIIPSPDLDAVECHRGCEQAGVLFPGHGHNRNGTGHNGSTHGPYRCS 186
Db      373 G-----PAPLTEL-----GLVRFVHA-----ALWGPE----- 396
Qy      187 SPHLVLSALTSNHNATYAFSGTHYWRDLTSRDGHSWPI--AHQWPGPSAVDAAFPSW 243
Db      397 -----KKKIYFFRGDYMRFHPTQRYVDN--FVPRRTDWRGVPSEIDAAF-- 440
Qy      244 BEKLYVQGTQVYVFLTKG 262
Db      441 ---QDADGVAYFLRG 452

```

Search completed: December 16, 2003, 06:53:08
Job time : 38 secs

[illegible]

C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #ext_change 10-Jul-1998

C/Accession: A44399

R/Label: O.; Wolf, C.; Limacher, J.M.; Hutin, P.; Wendling, C.; Leheur, M.; Basset, P.

J. Cell Biol. 119, 997-1002, 1992

A>Title: The breast cancer-associated stromelysin-3 gene is expressed during mouse mammary

A/Reference number: A44399; MUID:93054930; PMID:1429845

A/Accession: A44399

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: nucleic acid

A/Residues: 1-492 <LRF>

C/Note: sequence extracted from NCBI backbone (NCBIP:117216)

C/Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloproteinase

C/Keywords: hydrolyase; metalloproteinase; zinc; zymogen

F:52-262/Domain: matrix metalloproteinase homology <MMP>

F:292-484/Domain: hemopexin repeat homology <HPN>

F:84-219,223,229/Binding site: zinc, catalytic (His, His, His, His) (inhibited) #status predicted

F:219,223,229/Binding site: zinc, catalytic (His) (active) #status predicted

F:220/Active site: Glu #status predicted

Query Match 8.3%; Score 181.5; DB 2; Length 492;

Best Local Similarity 25.2%; Pred. No. 2.4e-07;

Matches 69; Conservative 26; Mismatches 86; Indels 93; Gaps 13;

Db 270 SPAPLSSQAGTDTNEILLEPPEPVCET----SPDVS-TTIGELFFKAGFVWRL 323

24 TPLPPTSAH-----GNVAGETKPPDVTYERSCDSGSPATTLDDNGTMLFFKGEFVWKS 78

79 HKWDR-----LISERKKNFSPVDAAFRGHSNVEFLIGDKVWVYPPKKKGYPKLLQ 133

Db 324 RSGRLQPEYRALASHRMQGLPSPVDAAFDAGQGIFFQAGQYWVYDGEKPYVLGAPLPSK 383

Qy 134 DEFGIPEPLDAVNECHRGSCQAEGLVFFQGHGHRNGTGHGNSTHGPEYWRCSPHVLVS 193

Db 384 LGLQG--SPVHAALV-----WGPE----- 400

Qy 194 ALTSDNHGATYAFSGSTHYWRLD--TSR-DGWSMPRIAH--WPGSPAVDAAFSWEKL 247

Db 401 -----KKKIYFPRGDIYWRFRPRTORVDN---PYPRSTDKRGVPSIEDAFA----- 444

Qy 248 YLVQSTQVYVFLTKG-----GYTVSGYK 272

Db 445 ---QDAEGYAFRLRHLWYKRPVKKVYKLEGFPR 475

RESULT 14

JC6197

stromelysin 3 (EC 3.4.24.-) - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 11-Apr-1997 #sequence_revision 09-May-1997 #ext_change 17-Mar-1999

C/Accession: JC6197

R/Okada, A.; Saez, S.; Mismun, Y.; Basset, P.

A>Title: Rat stromelysin 3: cDNA cloning from healing skin wound, activation by furin and

A/Reference number: JC6197; MUID:97208872; PMID:9055814

A/Contents: Skin wounds

A/Accession: JC6197

A/Molecule type: mRNA

A/Residues: 1-491 <OKA>

A/Cross-references: GB:U46034

C/Comment: This protein is a member of the matrix metalloproteinase family.

C/Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloproteinase

C/Keywords: hydrolyase; metalloproteinase; zinc; zymogen

F:52-261/Domain: matrix metalloproteinase homology <MMP>

F:291-483/Domain: hemopexin repeat homology <HPN>

F:84-218,222,228/Binding site: zinc, catalytic (His, His, His, His) (inhibited) #status predicted

F:218,222,228/Binding site: zinc, catalytic (His) (active) #status predicted

F:219/Active site: Glu #status predicted

Query Match 8.3%; Score 180.5; DB 2; Length 491;

Best Local Similarity 26.7%; Pred. No. 2.8e-07;

Matches 68; Conservative 22; Mismatches 84; Indels 81; Gaps 13;

Db 24 TPLPPTSAHGNVASETK---PDPEV-TERSGDSGSPATTLDDNGTMLFFKGEFVWKS 78

C:Species: Homo sapiens (man)
C:Date: 13-Aug-1996 #sequence revision 30-Sep-1992 #text change 08-Dec-2000
A:Accession: A37308, S22766, 157620, A00996, D29157, A44518, S06133, B60964, S10595, S533
R:Template: N.S., Brown, P.D., Levy, A.T., Margulies, I.M.K., Liotta, L.A., Stetler-Stevenson
Cancer Res. 50, 5431-5437, 1990
A:Title: Cloning and characterization of human tumor cell interstitial collagenase.
A:Reference number: A37308; PMID:90352587; PMID:2167156
A:Accession: A37308
A:Molecule type: mRNA
A:Residues: 1-469 <TM>
A:Cross-references: GB:R54925; NID:G30125; PIDN:CAA38691.1; PID:G30126
R:Brinckerhoff, C.E., Ruby, P., Austin, S.D., Fini, M.E., White, H.D.
J. Clin. Invest. 79, 542-546, 1987
A:Title: Molecular cloning of human synovial cell collagenase and selection of a single
A:Reference number: S22766; PMID:87109799; PMID:30277129
A:Accession: S22766
A:Molecule type: DNA
A:Residues: 1-63, 65-70

A:Cross-references: EMBL:M15996; NID:G180666; PIDN:AAA35700.1; PID:G180667
R:Angel, P., Baumann, I., Stein, B., Delius, H., Rahmsdorf, H.-J., Herrlich, P.
Mol. Cell. Biol. 7, 2256-2266, 1987
A:Title: 12-O-tetradecanoyl-phorbol-13-acetate induction of the human collagenase gene 1
A:Reference number: 157620; PMID:87257941; PMID:3037355
A:Accession: 157620
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-35 <RS>
A:Cross-references: GB:M16567; NID:G180668; PIDN:AAA52033.1; PID:G180669
R:Goldberg, G.I., Wilhelm, S.M., Kronberger, A., Bauer, E.A., Grant, G.A., Eisen, A.Z.
J. Biol. Chem. 261, 6600-6605, 1986
A:Title: Human fibroblast collagenase. Complete primary structure and homology to an onc
A:Reference number: A00996; PMID:86196089; PMID:3009463
A:Accession: A00996
A:Molecule type: mRNA
A:Residues: 1-114, 'R', 116-409, 'S', 411-469 <GOL>
A:Cross-references: GB:M13509; NID:G180664; PIDN:AAA55699.1; PID:G180665
A:Note: part of this sequence was confirmed by protein sequencing of the proenzyme
R:Whitham, S.E., Murphy, G., Angel, P., Rahmsdorf, H.-J., Smith, B.-J., Lyons, A., Harris,
Biochem. J. 240, 913-916, 1986
A:Title: Completion of human stromelysin and collagenase by cloning and sequence analysis
A:Reference number: A90336; PMID:87156645; PMID:3030290
A:Accession: D29157
A:Molecule type: mRNA
A:Residues: 1-199, 'H', 201-207, 'T', 209-469 <WHI>
A:Cross-references: EMBL:X05231; NID:G38266; PIDN:CAA28858.1; PID:G38267
A:Note: parts of this sequence, including the amino end of the proenzyme and of the metua
R:Birkedal-Hansen, B., Moore, W.G.I., Taylor, R.E., Bhown, A.S., Birkedal-Hansen, H.
Biochemistry 27, 6751-6758, 1988
A:Title: Monoclonal antibodies to human fibroblast procollagenase. Inhibition of enzymatic
end of the activated enzyme.
A:Reference number: A44518; PMID:89062403; PMID:2461732
A:Accession: A44518
A:Molecule type: protein
A:Residues: 270-305 <BIR>
R:Clark, I.M., Cawston, T.E.
Biochem. J. 263, 201-206, 1989
A:Title: Fragments of human fibroblast collagenase. Purification and characterization.
A:Reference number: S06132; PMID:90104231; PMID:2557822
A:Accession: S06132
A:Status: preliminary
A:Molecule type: protein
A:Residues: 100-102, 'P', 104-107, 'XX', 110-112, 270-277, 'X', 279-280, 'X', 282-287 <CLA>
R:Lark, M.W., Malachuk, L.A., Shah, T.K., Vanmiddlesworth, J., Cameron, P.M., Linn, T.Y.
Connect. Tissue Res. 25, 49-65, 1990
A:Title: Production and purification of prostromelysin and procollagenase from IL-1 beta
A:Reference number: A60964; PMID:91059606; PMID:2173990
A:Accession: B60964
A:Molecule type: protein
A:Residues: 24-35, 100-108, 270-272, 'X', 274, 'X', 276 <LAR>
R:Suzuki, K., Nagase, H., Ito, A., Enghild, J.J., Salvesen, G.
Biol. Chem. Hoppe-Seyler 371(Suppl.), 305-310, 1990
A:Title: The role of matrix metalloproteinase 3 in the stepwise activation of human rheum
A:Reference number: S10595; PMID:90380300; PMID:2169257

QY 195 LTSDNHGATYAFSGTHYWRDLTSR---DGMHWPJAHQPOGSAVDAFAFSEKLYLVQ 251
 Db 399 -----TYFFVGDKYKRFDETRHVMQKGFPRQITDQFPGIEPQVAVLHFFGFFYFR 450
 QY 252 GTQYVVF 258
 Db 451 GSSQFEF 457

RESULT 8

KCHS2
 scromelysin 2 (EC 3.4.24.22) precursor [validated] - human
 N:Alternate names: matrix metalloproteinase 10 (MMP10); transin-2
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 08-Dec-2000
 C:Accession: A28816; A47496
 R:Miller, D.; Quantin, B.; Gesnel, M.C.; Millon-Collard, R.; Abecassis, J.; Breathnach, Biochem. J. 253, 187-192, 1988
 A:Title: The collagenase gene family in humans consists of at least four members.
 A:Reference number: A90339; MUID:88339885; PMID:2844164
 A:Accession: A28816
 A:Molecule type: mRNA
 A:Residues: 1-476 <MUI>
 A:Cross-references: EMBL:X07820; NID:936628; PIDN:CAA30679.1; PID:936629
 A>Note: mRNA for this protein was detected in several human tumors
 R:Mindor, L.U.; Grenett, H.; Birkedal-Hansen, B.; Boddén, M.K.; Engler, J.A.; Birkedal-J. Biol. Chem. 268, 17341-17347, 1993
 A:Title: Cell type-specific regulation of SL-1 and SL-2 genes. Induction of the SL-2 gene
 A:Reference number: A47496; MUID:93352520; PMID:8349617
 A:Accession: A47496
 A:Molecule type: protein
 A:Residues: 17-33 <MIN>
 C:Comment: This enzyme degrades various extracellular matrix proteins, including fibronectin;
 C:Genetics:
 A:Gene: GDB:MMP10; STMY2
 A:Cross-references: GDB:120392; OMIM:185260
 A:Map position: 11q22.3-11q23
 C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotease
 C:Keywords: calcium; extracellular matrix; fibroblast; glycoprotein; hydrolase; metalloprotein; signal sequence #status predicted <SIG>
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-476/Product: prostromelysin 2 #status experimental <PRO>
 F:17-98/Domain: activation peptide #status predicted <ACT>
 F:59-263/Domain: matrix metalloproteinase homology <MMP>
 F:89-96/Region: autoinhibitory
 F:99-476/Product: stromelysin 2 #status predicted <MAT>
 F:283-476/Domain: hemopexin repeat homology <PXN>
 F:91,217,221,227/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status
 F:119/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:217,221,227/Binding site: zinc, catalytic (His) (active) #status predicted
 F:218/Active site: Glu #status predicted
 F:289-476/Disulfide bonds: #status predicted

Query Match 8.7%; Score 189.5; DB 1; Length 476;
 Best Local Similarity 28.4%; Pred. No. 4.8e-08;
 Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

QY 25 PLPTSAHGNAVAGETKDPDVTERTCSQWSPDA-TLTDNDGTMFFKGEVVMKSHKDR 83
 Db 272 PLVLPK---SVPSGEMP-----AAYCDPALSPDALSTL-RGEYLFKDRFMRSHNP 321
 QY 84 E-----LTSERKNPSPVDAFAFR-QGHSVFLIKGDKVWV-----YP 120
 Db 322 EPEPHLISAFWPSLPSYIDAAVEVNSRDTVFIFKGNEMWARGNEVQAGYRGHITLGF 381
 QY 121 P-----EKK-----EKGYPKLLQDEPFGIPLDAVE 148
 Db 382 PTKRKDAVSGDEKKKTYFFPAADKYWRFDENSGMEGPFRLADDPGVEPKVDATL- 440
 QY 149 CHRGECAQAGVLPF 162
 Db 441 -----QAFGFYF 448

RESULT 9

KCHS2
 scromelysin 2 (EC 3.4.24.22) precursor - rat
 N:Alternate names: matrix metalloproteinase 10 (MMP10); transin-2
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C:Accession: B26403; A41775; S26498
 R:Breathnach, R.; Marislian, L.M.; Gesnel, M.C.; Steud, A.; Leroy, P. Nucleic Acids Res. 15, 1139-1151, 1987
 A:Title: Sequences coding for part of oncogene-induced transin are highly conserved in a
 A:Reference number: A26403; MUID:87146421; PMID:3547333
 A:Accession: B26403
 A:Molecule type: mRNA
 A:Residues: 1-476 <BRE>
 A:Cross-references: EMBL:X05083; NID:957388; PIDN:CAA28739.1; PID:957389
 A>Note: intron positions were determined by comparison of the cDNA sequence to genomic
 A>Note: mRNA for this protein was expressed in several transformed rat embryo fibroblast
 R:Chan, J.C.; Scanlon, M.; Zhang, H.Z.; Jia, L.B.; Yu, D.H.; Hung, M.C.; French, M.; Eas J. Biol. Chem. 267, 1099-1103, 1992
 A:Title: Molecular cloning and characterization of v-mos-activated transformation-associ
 A:Reference number: A41775; MUID:92112748; PMID:1370458
 A:Accession: A41775
 A:Molecule type: mRNA
 A:Residues: 1-476 <CRA>
 A:Cross-references: GB:M65253; NID:9207150; PIDN:AAA42202.1; PID:9207151
 A>Note: sequence extracted from NCBI backbone (NCBIP:76184)
 R:de Vonghe, M.W.; Mukherjee, B.B.
 A:Oncogene 7, 109-119, 1992
 A:Title: Transformation of normal rat kidney cells by v-K-ras enhances expression of tra
 A:Reference number: S26498; MUID:92158347; PMID:1741158
 A:Accession: S26498
 A:Molecule type: protein
 A:Status: preliminary; translation not shown
 A:Residues: 31-103, 'L', 241-242, 'TQMEKRP', 251, 'L', 253-254, 'CE', 293-294, 'L', 296 <DEV>
 A:Cross-references: EMBL:X64020
 C:Genetics:
 A:Introns: 35/3; 117/2; 167/1; 209/1; 264/1; 311/2; 356/1; 409/2; 444/1
 C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotease
 C:Keywords: calcium; extracellular matrix; fibroblast; glycoprotein; hydrolase; metalloprotein; signal sequence #status predicted <SIG>
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-476/Product: prostromelysin 2 #status predicted <PRO>
 F:18-99/Domain: activation peptide #status predicted <ACT>
 F:60-264/Domain: matrix metalloproteinase homology <MMP>
 F:90-97/Region: autoinhibitory
 F:100-476/Product: stromelysin 2 #status predicted <MAT>
 F:283-476/Domain: hemopexin repeat homology <PXN>
 F:92,218,222,228/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status
 F:120/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:218,222,228/Binding site: zinc, catalytic (His) (active) #status predicted
 F:219/Active site: Glu #status predicted
 F:289-476/Disulfide bonds: #status predicted

Query Match 8.6%; Score 187; DB 1; Length 476;
 Best Local Similarity 26.4%; Pred. No. 7.8e-08;
 Matches 52; Conservative 23; Mismatches 48; Indels 74; Gaps 8;

QY 21 ALAPPLTSAHGNAVAGETKDPDVTERTCSQWSPDA-TLTDNDGTMFFKGEVVMKSHK 80
 Db 271 ATVPVPSVS-----PKETPVKCDPALSPFAVTV-LRGEFLFFKDRHFMRTQ 318
 QY 81 WDRE-----LTSERKNPSPVDAFAFRQGH-NSVFLIKGDKVWV----- 118
 Db 319 WNEPEPHLISAFWPSLPSGIDAAVEANNDKRVLIIFKGSQWMAARGNEVQAGYRKRIHTL 378
 QY 119 -YPP-----EKK-----EKGYPKLLQDEPFGIPLDA 145
 Db 379 GFPEPTVKKIDAAVEKEKKKTYFFVGDKYWRFDTRQLMDGPFRLITDQPGIEPQVDA 438
 QY 146 AVECHRGCAQAGVLPF 162
 Db 439 VL-----HAFGFYF 448

A:Title: Identification of an interleukin-6 responsive element and characterization of h
A:Reference number: A43079; MUID:92287130; PMID:1599480
A:Accession: A43079
A:Molecule type: DNA
A:Residues: 1-14 <NAG>
A:Cross-references: GB:X60006; NID:g416183; PIDN:CAA42621.1; PID:g416184
R:Niklila, H.; Gellin, J.D.; Muller-Eberhard, U.
Biochemistry 30, 823-829, 1991
A:Title: Rat hemopexin. Molecular cloning, primary structural characterization, and anal
A:Reference number: A38399; MUID:91105180; PMID:1988069
A:Accession: A38399
A:Molecule type: mRNA
A:Residues: 15-460 <NIK>
A:Cross-references: GB:J05306
R:Weller, D.; Cheng, K.C.; Muller-Eberhard, U.
Biochem. Biophys. Res. Commun. 155, 622-625, 1988
A:Title: N-terminal amino acid sequences of the hemopexins from chicken, rat and rabbit.
A:Reference number: A90148; MUID:88339942; PMID:3421961
A:Accession: A90148
A:Molecule type: Protein
A:Residues: 24-37, 'C', '39-48', 'KW', '51', 'X', '53' <MEL>
R:Swerts, J.P.; Soula, C.; Sagoc, Y.; Guinaudy, M.J.; Guillemot, J.C.; Ferrara, P.; Dupr
J. Biol. Chem. 267, 10596-10600, 1992
A:Title: Hemopexin is synthesized in peripheral nerves but not in central nervous system
A:Reference number: A38139; MUID:92268104; PMID:1587840
A:Accession: A38139
A:Molecule type: Protein
A:Residues: 24-35, 'X', '37', 'X', '39-43' <SME>
A:Experimental source: sciatic nerve; skeletal muscle
A:Note: the amino-terminal sequence of the mature protein was determined
C:Comment: Hemopexin is a serum glycoprotein that binds heme and transports it to the li
C:Superfamily: hemopexin; hemopexin repeat homology
C:Keywords: acute phase; chromoprotein; duplication; glycoprotein; heme; iron; metallopr
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-460/Product: hemopexin #status predicted <HPX>
F:44-230/Domain: hemopexin repeat homology <PX1>
F:250-488/Domain: hemopexin repeat homology <PX2>
F:38,64,186,240,246/Binding site: carboxylate (Asn) (covalent) #status predicted
F:50-230,148-153,187-199,255-458,364-406,416-433/Diulfide bonds: #status predicted
F:79,149/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 71.2%; Score 1556.5; DB 1; Length 460;
Best Local Similarity 63.8%; Pred. No. 9.1e-123;
Matches 296; Conservative 35; Mismatches 54; Indels 79; Gaps 6;

Qy 1 MARVLGAPVALGLMSLCSLAIAIPLPTSAHGNVAE--TKPDPTVTERCSGMSFDA 58
Db 1 MARVLGAPVALGLMSLCSLAIAIPLPTSAHGNVAE--TKPDPTVTERCSGMSFDA 58
Qy 59 TLLDNGTMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 118
Db 59 TLLDNGTMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 118
Qy 59 TTHDNGTMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 117
Db 59 TTHDNGTMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 117
Qy 119 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 165
Db 119 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 165
Qy 118 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 177
Db 118 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 177
Qy 166 -----GH-- 167
Db 166 -----GH-- 167
Qy 178 ERSKPAVGNCTAALRMLEERYCFQGNKFLRNPVTEGVPPRYPLDARDYFISCRGK 237
Db 178 ERSKPAVGNCTAALRMLEERYCFQGNKFLRNPVTEGVPPRYPLDARDYFISCRGK 237
Qy 168 -RNGTGHGNTHTHGPVYRCSPHLVLSALTSNDNGATYAFSGTHYMRLDTSRDGMHSMPI 226
Db 168 -RNGTGHGNTHTHGPVYRCSPHLVLSALTSNDNGATYAFSGTHYMRLDTSRDGMHSMPI 226
Qy 238 LRNGTAGNNTHTH--PMISRCNADPGLSLSDHNGATYAFSGTHYMRLDTSRDGMHSMPI 295
Db 238 LRNGTAGNNTHTH--PMISRCNADPGLSLSDHNGATYAFSGTHYMRLDTSRDGMHSMPI 295
Qy 227 AHQWPGPSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIIL 286
Db 227 AHQWPGPSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIIL 286
Qy 296 AHHPQGPSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIIL 355
Db 296 AHHPQGPSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIIL 355
Qy 287 DSUVAATCPSSSLHTHIMAGRLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNS 346
Db 287 DSUVAATCPSSSLHTHIMAGRLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNS 346
Qy 356 DTIDAAFSRCSKSLYVTSGRRLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNS 415
Db 356 DTIDAAFSRCSKSLYVTSGRRLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNS 415

Qy 347 CSANGPGILYHIGPNLYCYSDVEKLNAAKALPQPONTSLIGCT 390
Db 416 CSANGPGILYHIGPNLYCYSDVEKLNAAKALPQPONTSLIGCT 390
Qy 416 CSANGPGILYHIGPNLYCYSDVEKLNAAKALPQPONTSLIGCT 390
Db 416 CSANGPGILYHIGPNLYCYSDVEKLNAAKALPQPONTSLIGCT 390

RESULT 4
A55486
hyaluronidase - pig
N:Alternate names: hemopexin
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Aug-1999
C:Accession: A55486
R:Zhu, L.; Hope, T.J.; Hall, J.; Davies, A.; Stern, M.; Muller-Eberhard, U.; Stern, R.;
J. Biol. Chem. 269, 32092-32097, 1994
A:Title: Molecular cloning of a mammalian hyaluronidase reveals identity with hemopexin.
A:Reference number: A55486; MUID:95096047; PMID:7798203
A:Accession: A55486
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-459 <ZHU>
A:Cross-references: GB:U14751; NID:g541627; PIDN:AAC48457.1; PID:g541628
C:Superfamily: hemopexin; hemopexin repeat homology
F:46-232/Domain: hemopexin repeat homology <PX1>
F:245-453/Domain: hemopexin repeat homology <PX2>

Query Match 70.7%; Score 1544; DB 2; Length 459;
Best Local Similarity 63.4%; Pred. No. 1e-121;
Matches 289; Conservative 33; Mismatches 64; Indels 70; Gaps 5;

Qy 1 MARVLGAPVALGLMSLCSLAIAIPLPTSAHGNVAE--TKPDPTVTERCSGMSFDA 58
Db 1 MARVLGAPVALGLMSLCSLAIAIPLPTSAHGNVAE--TKPDPTVTERCSGMSFDA 58
Qy 59 TLLDNGTMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 118
Db 59 TLLDNGTMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 118
Qy 61 STLDHGAHMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 120
Db 61 STLDHGAHMLFFKGEFVWKSHKMDRELISERKNFPPVDAAFROGHSVFLIKGDKYV 120
Qy 119 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 164
Db 119 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 164
Qy 121 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 179
Db 121 YPEKKEKGYPKLLQDEFPPIPSPLDAVACHRECOAEGVLFPOG----- 179
Qy 165 -----GH-- 173
Db 165 -----GH-- 173
Qy 180 ERLPAPVGNCSAMRWRISRYCFRGNQFLRFDPVTVGVDPKPRDVRDVFMSCPGRGHAH 239
Db 180 ERLPAPVGNCSAMRWRISRYCFRGNQFLRFDPVTVGVDPKPRDVRDVFMSCPGRGHAH 239
Qy 174 GNTHTHGPVYRCSPHLVLSALTSNDNGATYAFSGTHYMRLDTSRDGMHSMPIAHQWPG 233
Db 174 GNTHTHGPVYRCSPHLVLSALTSNDNGATYAFSGTHYMRLDTSRDGMHSMPIAHQWPG 233
Qy 240 RNATHRGDD--RCPDVLVLTALSDNNGATYAFSGTHYMRLDTSRDGMHSMPIAHQWPG 297
Db 240 RNATHRGDD--RCPDVLVLTALSDNNGATYAFSGTHYMRLDTSRDGMHSMPIAHQWPG 297
Qy 234 PSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIILDSVDAAF 293
Db 234 PSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIILDSVDAAF 293
Qy 298 PSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIILDSVDAAF 357
Db 298 PSAVDAAFSBEKLYLVQGTQVYVFLTKGGYTVLVSGYPKRLEKEVGTPHGIILDSVDAAF 357
Qy 294 ICPSSSLHTHIMAGRLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNSCSANGPG 353
Db 294 ICPSSSLHTHIMAGRLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNSCSANGPG 353
Qy 358 VCPGTSRLTHYHAGKLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNSCSANGPG 417
Db 358 VCPGTSRLTHYHAGKLMWLDLKSQAQWTWELPMPHKKVGDALCMESLGPNSCSANGPG 417
Qy 354 LYLIHGNLYCYSDVEKLNAAKALPQPONTSLIGCT 389
Db 354 LYLIHGNLYCYSDVEKLNAAKALPQPONTSLIGCT 389
Qy 418 LYTIHGNLYCYSDVEKLNAAKALPQPONTSLIGCT 453
Db 418 LYTIHGNLYCYSDVEKLNAAKALPQPONTSLIGCT 453

RESULT 5
I50485
Warm temperature acclimation-related 65-kDa protein, Wap65 - goldfish
C:Species: Carassius auratus (goldfish)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I50485
R:Kikuchi, K.; Yamashita, M.; Watabe, S.; Aida, K.
J. Biol. Chem. 270, 17087-17092, 1995
A:Title: The warm temperature acclimation-related 65-kDa protein, Wap65, in goldfish
A:Reference number: I50485; MUID:95340486; PMID:7615502
A:Accession: I50485

A:Residues: 24-49;62-67;178-192;236-250;443-457 <TA2>
 C:Comment: Hemopexin is a serum glycoprotein that binds heme and transports it to the 11
 C:Genetic:
 A:Gene: GDB:HPX
 A:Cross-references: GDB:120054; OMIM:142290
 A:Map position: 11p15.4-11p15.4
 A:Intons: 28/2; 48/1; 72/1; 112/3; 164/1; 235/1; 279/1; 322/3; 377/1
 C:Superfamily: hemopexin; hemopexin repeat homology
 C:Keywords: acute phase; duplication; glycoprotein; heme binding; plasma
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-462/Product: hemopexin #status experimental <HPX>
 F:44-231/Domain: hemopexin repeat homology <PX2>
 F:52-460/Domain: hemopexin repeat homology <PX2>
 F:50-231,149-154,188-200,257-460,418-435/Disulfide bonds: #status experimental
 F:54,187,240,246,453/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 97.9%; Score 2139.5; DB 1; Length 462;
 Best Local Similarity 84.6%; Pred. No. 1,1e-171;
 Matches 391; Conservative 0; Mismatches 0; Indels 71; Gaps 1;

QY 1 MARVLGAPVALGWSLCSLAITPLPPTSAHGNVAEGTKPPDPTTERCSDDGMSFDATT 60
 DB 1 MARVLGAPVALGWSLCSLAITPLPPTSAHGNVAEGTKPPDPTTERCSDDGMSFDATT 60
 QY 61 LDDNGTMLPFKEGFVWVSKHMDRELISERKWNFPSPVDAAFROGHNSVFLIKDKXWVY 120
 DB 61 LDDNGTMLPFKEGFVWVSKHMDRELISERKWNFPSPVDAAFROGHNSVFLIKDKXWVY 120
 QY 121 PEKKEKGYPKLLDDEFPGIPSPIDAAVECHRGECQAGVLFQGDREMFWDATGTMKER 163
 DB 121 PEKKEKGYPKLLDDEFPGIPSPIDAAVECHRGECQAGVLFQGDREMFWDATGTMKER 180
 QY 164 -----GHGHRN 169
 DB 181 SWEAVNGSSALRWLGRYYCFQGNQFLRDPVAVGEVPPRYRDVRYFMPGCRGHGHRN 240
 QY 170 GTGHGNSHTHGEPEYWCSPHLVLSALTSDNHGATYAFSGTHYWRDLTSDRGHMSWPIAHQ 229
 DB 241 GTGHGNSHTHGEPEYWCSPHLVLSALTSDNHGATYAFSGTHYWRDLTSDRGHMSWPIAHQ 300
 QY 230 WPGGSAVDAAFSWEKLYLVGCTQYVFLTKGGYTLVSGYPRLEKEVTPHGIILDSV 289
 DB 301 WPGGSAVDAAFSWEKLYLVGCTQYVFLTKGGYTLVSGYPRLEKEVTPHGIILDSV 360
 QY 290 DAAFLCPGSSRLHIMGRRLMWLDLKSQAQATWTELPMHEKVDGALCMEKSLGPNSSCA 349
 DB 361 DAAFLCPGSSRLHIMGRRLMWLDLKSQAQATWTELPMHEKVDGALCMEKSLGPNSSCA 420
 QY 350 NGPGLYLHGPMLYCYSDVEKLNAAKALPOPOVNTSLLGCTH 391
 DB 421 NGPGLYLHGPMLYCYSDVEKLNAAKALPOPOVNTSLLGCTH 462

RESULT 2
 OORR
 hemopexin precursor - rabbit
 N:Alternate names: beta-1b-glycoprotein
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 21-Sep-1993 #sequence, revision 13-Jan-1995 #text_change 03-Mar-2000
 C:Accession: A46006; A61426; A61426; B31514
 R:Morgan, W.T.; Muster, P.; Tatum, F.; Kao, S.M.; Alam, J.; Smith, A.
 J. Biol. Chem. 268, 6256-6262, 1993
 A:Title: Identification of the histidine residues of hemopexin that coordinate with heme
 A:Reference number: A46006; M01D:93203213; PMID:7681064
 A:Accession: A46006
 A:Molecule type: mRNA
 A:Residues: 1-459 <MOR>
 A:Note: sequence extracted from NCBI backbone (NCBIN:127918, NCBI:P.127919)
 A:Accession: B46006
 A:Molecule type: protein
 A:Residues: 26-33;104-108;239-242; 'X', 243-244 <MO2>
 A:Note: residues 42-46 are shown as 'TKPEA' in the Fig. 2 alignment, with no explanation

R:Muster, P.; Tatum, F.; Smith, A.; Morgan, W.T.
 J. Protein Chem. 10, 123-128, 1991
 A:Title: Further characterization of structural determinants of rabbit hemopexin function
 A:Reference number: A61426; M01D:91273754; PMID:2054057
 A:Accession: A61426
 A:Molecule type: protein
 A:Residues: 26-41; 'TKPEA', 47-51 <MUS>
 R:Wellner, D.; Cheng, K.C.; Muller-Eberhard, U.
 Biochem. Biophys. Res. Commun. 155, 622-625, 1988
 A:Title: N-terminal amino acid sequences of the hemopexins from chicken, rat and rabbit.
 A:Reference number: A90148; M01D:88339942; PMID:3421961
 A:Accession: B31514
 A:Molecule type: protein
 A:Residues: 26-33; 'X', 35-41; 'TKPEA', 47-51; 'W', 53 <MEU>
 C:Comment: Hemopexin is a serum glycoprotein that binds heme and transports it to the 11
 C:Keywords: The JEN-14 monoclonal antibody, which blocks the interaction of hemopexin with
 C:Superfamily: hemopexin; hemopexin repeat homology
 C:Keywords: acute phase; hemopexin; duplication; glycoprotein; heme; iron; metallopro
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-459/Product: hemopexin #status experimental <HPX>
 F:46-232/Domain: hemopexin repeat homology <PX1>
 F:249-457/Domain: hemopexin repeat homology <PX2>
 F:24/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:52-232,150-155,189-201,254-457,363-405,415-432/Disulfide bonds: #status predicted
 F:56,188,243/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:81,151/Binding site: heme iron (His) (axial ligand) #status experimental

Query Match 77.2%; Score 1686; DB 1; Length 459;
 Best Local Similarity 67.3%; Pred. No. 1,2e-133;
 Matches 311; Conservative 28; Mismatches 49; Indels 74; Gaps 4;

QY 1 MARVLGAPVALGWSLCSLAITPLPPTSAHGNVAEGTKPPDPTTERCSDDGMSFDPA 58
 DB 1 MKASGIPALGWSLCSLAITPLPPTSAHGNVAEGTKPPDPTTERCSDDGMSFDPA 60
 QY 59 TTLDDNGTMLPFKEGFVWVSKHMDRELISERKWNFPSPVDAAFROGHNSVFLIKDKXWV 118
 DB 61 TTLDDNGTMLPFKEGFVWVSKHMDRELISERKWNFPSPVDAAFROGHNSVFLIKDKXWV 119
 QY 119 YPEKKEKGYPKLLDDEFPGIPSPIDAAVECHRGECQAGVLFQGH-----165
 DB 120 YPEKKEKGYPKLLDDEFPGIPSPIDAAVECHRGECQAGVLFQGH-----165
 QY 166 -----GHHRN 169
 DB 180 EESWPAVNGCTALRWLGRYYCFQGNQFLRDPVAVGEVPPRYRDVRYFMPGCRGHGHRN 238
 QY 170 GTGHGNSHTHGEPEYWCSPHLVLSALTSDNHGATYAFSGTHYWRDLTSDRGHMSWPIAHQ 229
 DB 239 -SSHRNSTGHGHESTRCDPDLVLSAMVSDNHGATYAFSGTHYWRDLTSDRGHMSWPIAHQ 297
 QY 230 WPGGSAVDAAFSWEKLYLVGCTQYVFLTKGGYTLVSGYPRLEKEVTPHGIILDSV 289
 DB 298 WPGGSAVDAAFSWEKLYLVGCTQYVFLTKGGYTLVSGYPRLEKEVTPHGIILDSV 357
 QY 290 DAAFLCPGSSRLHIMGRRLMWLDLKSQAQATWTELPMHEKVDGALCMEKSLGPNSSCA 349
 DB 358 DAAFLCPGSSRLHIMGRRLMWLDLKSQAQATWTELPMHEKVDGALCMEKSLGPNSSCA 417
 QY 350 NGPGLYLHGPMLYCYSDVEKLNAAKALPOPOVNTSLLGCTH 391
 DB 418 SGPNLYLHGPMLYCYSDVEKLNAAKALPOPOVNTSLLGCTH 459

RESULT 3
 OORR
 hemopexin precursor - rat
 N:Alternate names: beta-1b-glycoprotein
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Jun-1991 #sequence, revision 13-Jan-1995 #text_change 03-Mar-2000
 C:Accession: A43073; A58399; A31514; A38139
 R:Nagae, Y.; Muller-Eberhard, U.
 Biochem. Biophys. Res. Commun. 185, 420-429, 1992

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 16, 2003, 05:47:17 ; Search time 34 Seconds
(without alignments)
1105.940 Million cell updates/sec

Title: US-09-900-448-2

Perfect score: 2185
Sequence: 1 MARVLGAPVALGLWSLMSL.....NAAKALPQPQNVTSILGCTH 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2139.5	97.9	462	1	OOBU hemopexin precursor
2	1686	77.2	459	1	OOBU hemopexin precursor
3	1556.5	71.2	460	1	COBT hemopexin precursor
4	1544	70.7	459	2	AS486 hyaluronidase - pi
5	447.5	20.5	445	2	Warm temperature a
6	202.5	9.3	470	2	A49499 metalloelastase HM
7	195.5	8.9	476	1	JC6505 stromelysin 2 (EC
8	189.5	8.7	476	1	KCHUS2 stromelysin 2 (EC
9	187	8.6	476	1	KCRBS2 stromelysin 2 (EC
10	184.5	8.4	468	1	KCRBI interstitial colla
11	183.5	8.4	467	1	KCHUN neutrophil collag
12	181.5	8.3	469	1	KCHUI interstitial colla
13	181.5	8.3	469	1	KCHUI interstitial colla
14	180.5	8.3	491	2	A44399 stromelysin 3 (EC
15	179.5	8.3	488	2	JC6197 stromelysin 3 (EC
16	177	8.1	477	1	S13423 stromelysin 3 (EC
17	177	8.1	475	1	KCRTH stromelysin 1 (EC
18	176	8.1	471	1	AS3711 stromelysin 1 (EC
19	173	7.9	477	1	KCMS51 collagenase 3 (EC
20	172	7.9	669	2	I38029 stromelysin 1 (EC
21	170.5	7.8	478	1	SGHUI matrix metalloprot
22	169.5	7.8	477	1	SGHUI vitronectin precu
23	169.5	7.8	521	2	I37252 stromelysin 3 (EC
24	167.5	7.7	508	2	JCS082 probable matrix me
25	166.5	7.6	466	2	A23685 matrix metalloprot
26	166	7.6	469	1	KCRPI interstitial colla
27	166	7.6	476	1	SGMSV interstitial colla
28	166	7.6	478	1	KCRBS1 vitronectin precu
29	158	7.2	459	2	JCS139 vitronectin precu

30	154.5	7.1	472	2	S29243 interstitial colla
31	151	6.9	469	1	KCBOI interstitial colla
32	150.5	6.9	582	2	I38028 matrix metalloprot
33	149.5	6.8	475	2	A38340 66k glycoprotein p
34	149.5	6.8	582	2	I84471 matrix metalloprot
35	147.5	6.8	462	2	A42401 matrix metalloprot
36	143.5	6.6	663	1	S46492 gelatinase A (EC 3
37	142.5	6.5	662	2	S70365 gelatinase A (EC 3
38	142	6.5	662	2	A42496 gelatinase A (EC 3
39	142	6.5	662	2	S34780 gelatinase A (EC 3
40	140.5	6.4	660	1	A28153 gelatinase A (EC 3
41	139	6.4	484	2	A40774 phosphocholine-bin
42	135	6.2	707	1	A34458 gelatinase B (EC 3
43	133.5	6.1	582	2	I48673 matrix metalloprot
44	124.5	5.7	616	2	JC7776 matrix metalloprot
45	124	5.7	483	2	JCS743 matrix metalloprot

ALIGNMENTS

RESULT 1
hemopexin precursor [validated] - human
N:Alternate names: beta-1B-glycoprotein
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence revision 02-Jul-1996 #text change 08-Dec-2000
C:Accession: I56456; I54212; A93566; A94069; A93328; A43791; A03263
R:Altruda, F.; Poli, V.; Restagno, G.; Silengo, L.
J. Mol. Evol. 27, 102-108, 1988
A>Title: Structure of the human hemopexin gene and evidence for intron-mediated evolution
A:Reference number: I56456; MUID:88316972; PMID:2842511
A:Accession: I56456
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-462 <RES>
A:Cross-references: GB:M36803; NID:G184495; PIDN:AAA58678.1; PID:G184497
R:Law, M.L.; Cai, G.Y.; Hartz, J.A.; Jones, C.; Kao, F.T.
Genomics 3, 48-52, 1988
A>Title: The hemopexin gene maps to the same location as the beta-globin gene cluster on
A:Reference number: I54212; MUID:89122012; PMID:3220477
A:Accession: I54212
A>Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 2-462 <RES>
A:Cross-references: GB:J03048; NID:G184487; PIDN:AAA52704.1; PID:G386789
R:Altruda, F.; Poli, V.; Restagno, G.; Argos, P.; Cortese, R.; Silengo, L.
Nucleic Acids Res. 13, 3841-3859, 1985
A>Title: The primary structure of human hemopexin deduced from cDNA sequence: evidence f
A:Reference number: A93566; MUID:85242073; PMID:2969777
A:Accession: A93566
A:Molecule type: mRNA
A:Residues: 22-462 <ALT>
R:Takahashi, N.; Takahashi, Y.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 73-77, 1985
A>Title: Complete amino acid sequence of human hemopexin, the heme-binding protein of se
A:Reference number: A94069; MUID:85113173; PMID:3855550
A:Accession: A94069
A:Molecule type: protein
A:Residues: 24-462 <TAK>
R:Frankikova, V.; Borvayk, U.; Klun, I.; Moravsek, L.
FEBS Lett. 178, 213-216, 1984
A>Title: Amino acid sequence of the N-terminal region of human hemopexin.
A:Reference number: A91328; MUID:85076955; PMID:6510521
A:Accession: A91328
A:Molecule type: protein
A:Residues: 24-255 <FRA>
R:Takahashi, N.; Takahashi, Y.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 2021-2025, 1984
A>Title: Structure of human hemopexin: O-glycosyl and N-glycosyl sites and unusual clust
A:Reference number: A43791; MUID:84193947; PMID:6371807
A:Accession: A43791
A:Molecule type: protein

DR PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
KM Hydrolase; Metalloprotease; zinc; Zymogen; Calcium; Signal; 3D-structure.
NM Collagen degradation; Extracellular matrix; Signal; 3D-structure.
FT STGNL 1 35
FT PROPEP 36 101 POTENTIAL.
FT CHAIN 102 492 ACTIVATION PEPTIDE (BY SIMILARITY).
FT DOMAIN 295 487 STROMELYSIN-3.
FT SITE 84 84 HEMOPLEXIN-LIKE.
FT METAL 219 219 CYSTEINE SWITCH (BY SIMILARITY).
FT ACT SITE 220 220 ZINC (CATALYTIC).
FT METAL 223 223 ZINC (CATALYTIC).
FT METAL 229 229 ZINC (CATALYTIC).
FT DISULFID 298 484 ZINC (CATALYTIC).
FT STRAND 102 105 BY SIMILARITY.
FT TURN 106 107
FT STRAND 108 111
FT STRAND 113 118
FT HELIX 127 142
FT TURN 143 144
FT STRAND 148 151
FT STRAND 159 164
FT STRAND 182 185
FT TURN 188 189
FT STRAND 193 198
FT TURN 199 200
FT STRAND 203 204
FT TURN 208 209
FT STRAND 211 212
FT HELIX 213 224
FT TURN 225 226
FT TURN 233 234
FT TURN 236 237
FT HELIX 251 261
SQ SEQUENCE 492 AA; 55441 MW; B54E260E4AB3D7C3 CRC64;

Query Match 8.44; Score 184.5; DB 1; Length 492;
Best Local Similarity 25.24; Pred. No. 2.9e-08;
Matches 70; Conservative 27; Mismatches 88; Indels 93; Gaps 13;

QY 20 LAIATPIPTSAH-----GNVAGETKPDVDYTERCSDGMSFDATTIDNGTMLFFRGER 74
DB 266 MAPTSPAPVLTSSQAGTDTNEIALLEPETPPVCET-----SFDVVS-TIRGELFFRQGF 319
QY 75 VMKSHKWDRE-----LISERKNFPPSPVDAAFRQHNVSFLIKGDKVWVYPPEKEKGYR 129
DB 320 VMRLRSGRLOPGYPALASRMQGLPSPYDAAFEDAQOIWFFQCAQYVYDGEKPVLGPA 379
QY 130 KLIQDEFPGIIPSPIDAAVECHRGECQAEVLPFOGHGRNGTGHGNSTHHGPEYMRCSPI 189
DB 380 PLKSLGIQG--SPVHALV-----WGPR----- 400
QY 190 LVLSALTSMDHGAITYASGTHYWRLD--TSR-DGMHSWPIAHQ---WPOGSAVDAAFSW 243
DB 401 -----KNKIYFRGGDYWRFHPRTORVDN---PVRSTDRGVPSSEIDAF-- 444
QY 244 EEKLYLVQGTQVYVFLTKG-----GYTLVSGYRK 272
DB 445 -----QDAGYAVFLRGHLVWKFDPVKVLEGFPR 475

Search completed: December 16, 2003, 06:48:52
Job time : 30 secs

CC -1- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN
 CC THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X.
 CC -1- CATALYTIC ACTIVITY: Cleaves preferentially one bond in native
 CC collagen. Cleavage of the triple helix of collagen at about three-
 CC quarters of the length of the molecule from the N-terminus, at
 CC 775-Gly-Ile-776 in the alpha-1(I) chain. Cleaves synthetic
 CC substrates and alpha-macroglobulins at bonds where Pl¹ is a
 CC hydrophobic residue.
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE
 CC ACTIVATION PEPTIDE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
 CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M17823; AAB88016.1; -
 DR EMBL; M17820; AAB88016.1; JOINED.
 DR EMBL; M17821; AAB88016.1; JOINED.
 DR EMBL; M17822; AAB88016.1; JOINED.
 DR EMBL; M19240; AAB88016.1; JOINED.
 DR EMBL; M25663; AAA31203.1; -
 DR PIR; A27500; KCRBI.
 DR HSSP; P03956; ICGI.
 DR MEROPS; M10.001; -
 DR InterPro; IPR000585; Hemopexin.
 DR InterPro; IPR001818; Matrxin.
 DR InterPro; IPR006026; Nzn_MTPeptide.
 DR InterPro; IPR006025; Zn_MTPeptide.
 DR Pfam; PF00045; hemopexin; 4.
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR Pfam; PF03933; Peptidase_M10_N; 1.
 DR PRINTS; PR00138; MATRIXIN.
 DR SMART; SM00120; HX; 4.
 DR SMART; SM00235; Zmnc; 1.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KM Hydroxylase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 KM Collagen degradation; Extracellular matrix; Signal.
 FT STGNAL 1 18
 FT PROPEP 19 98
 FT CHAIN 99 468
 FT DOMAIN 274 468
 FT SITE 91 91
 FT METAL 217 217
 FT ACT_SITE 218 218
 FT METAL 221 221
 FT METAL 227 227
 FT CARBOHYD 119 119
 FT DISULFID 277 465
 SQ SEQUENCE 468 AA; 53739 MW; DA90538919952B8C CRC64;

Query Match 8.4%; Score 184.5; DB 1; Length 468;
 Best Local Similarity 24.1%; Pred. No. 2.7e-08;
 Matches 54; Conservative 33; Mismatches 80; Indels 57; Gaps 7;

QY 44 PDVTERSDGKSPFA--TTLDNGTMTLFFKGFVWKSXKHWDEL-----ISEWKNKPPSPVD 98
 Db 271 PDPKVCDSKLTPTAITT--RGSIMPFKDFYKRNAPYSEVLENFISVFWPLPGLQ 328
 QY 99 AAFROGH-NSVFLTKGDKVWVYPPKKKGYKLLQDEFGIPGLDAVAVECHGEGCAE 157
 Db 329 AAYVVAHDELTFKGNKRYVVOGQNELPGYKIHSSF----- 367
 QY 158 GVLFQGHGHNIGTGHNSSTHGEYKRCSPHLVLSALTSDNHGATYAFSGTHWRLD-- 215

Db 368 -----GPRFSVNH-----IDAASEDGTGTYFFVANKRYRDEY 402
 QY 216 -TSRDGWSHPIAHWQPGSAVDAAPSWEKLYVGOTGVYV 258
 Db 403 KRSDIAGYPMIEYDFPGIGNKVDVAVPKDGFYFFHGTROYKF 446

RESULT 15
 ID MML1_MOUSE STANDARD; PRT; 492 AA.
 AC 002853;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Stromelysin-3 precursor (EC 3.4.24.-) (Matrix metalloproteinase-11)
 DE (MMP-11) (ST) (SL-3).
 GN MMP11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93054930; Pubmed=1429845;
 RA Lefebvre O., Wolf C., Limacher J.M., Rutin P., Wendling C.,
 RA Lemer M., Bassot P., Rio M.C.;
 RT "The breast cancer-associated stromelysin-3 gene is expressed during
 RT mouse mammary gland apoptosis";
 RT J. Cell Biol. 119:997-1002 (1992).
 RN [2]
 RP REVISIONS.
 RA Lefebvre O.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 102-265.
 RX MEDLINE=21154142; Pubmed=11254383;
 RA Gall A.-L., Ruff M., Kannan R., Cunnase P., Viotakis A., Dive V.,
 RA Rio M.-C., Bassot P., Moras D.;
 RT "Crystal structure of the stromelysin-3 (MMP-11) catalytic domain
 RT complexed with a phosphinic inhibitor mimicking the
 RT transition-state";
 RT J. Mol. Biol. 307:577-586 (2001).
 RL J. Mol. Biol. 307:577-586 (2001).
 CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE PROGRESSION OF
 CC EPITHELIAL MALIGNANCIES.
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN THE MAMMARY GLAND
 CC DURING APOPTOSIS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
 CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z12604; CA78248.1; -
 DR PDB; 1HVS; 28-MAR-01.
 DR MEROPS; M10.007; -
 DR MGD; MGI:97008; MMP11.
 DR InterPro; IPR000585; Hemopexin.
 DR InterPro; IPR001818; Matrxin.
 DR InterPro; IPR006026; Nzn_MTPeptide.
 DR InterPro; IPR006025; Zn_MTPeptide.
 DR Pfam; PF00045; hemopexin; 4.
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR PRINTS; PR00138; MATRIXIN.
 DR SMART; SM00120; HX; 4.
 DR SMART; SM00235; Zmnc; 1.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.

```

Db 272 PLVPTK---SVPSGSEMP-----AKCDPALSPDAISTL--RGXYLFKDKYFRRRSHWNP 321
Qy 84 E-----LISERWKVPSPVDAAFR-QGHNSVFLIKGDKVWV-----YF 120
Db 322 EPEFHLSAIFWPSYLDAAVEVNSRDYFIFKGFEMFMRNGENVOAGYPRGIIHTLGF 361
Qy 121 P-----EKK-----EKGYPKLLDPEFGIPSPDLAAVE 148
Db 382 PIRKIDAAVSDKEKKTYFPADKWRFPDENSQSGMEQSPRLIADDFEVEKVDVAVL- 440
Qy 149 CHRGECAEGVLFP 162
Db 441 -----QAFGFYF 448

RESULT 13
MM10 RAT STANDARD; PRT; 476 AA.
AC P07152;
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Stromelysin-2 precursor (EC 3.4.24.22) (Matrix metalloproteinase-10)
GN (MMP-10) (Transin-2) (SL-2) (Transformation-associated protein 34A).
MP10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8716421; PubMed=3547333;
RA Breachmach R., Matriasian L.M., Gesnel M.-C., Staud A., Leroy P.;
RT "Sequences coding for part of oncogene-induced transin are highly
RT conserved in a related rat gene."
RL Nucleic Acids Res. 15:1139-1151(1987).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112748; PubMed=1370458;
RA Chan J.C., Scanlon M., Zhang H.Z., Jia L.B., Yu D., Hung M.C.,
RA French M., Eastman E.M.;
RT "Molecular cloning and characterization of v-mos-activated
RT transformation-associated proteins."
RL J. Biol. Chem. 267:1099-1103(1992).
CC - FUNCTION: CAN DEGRADE FIBRONECTIN, GELATIN OF TYPE I, III, IV,
CC AND V; WEAKLY COLLAGENS III, IV, AND V. ACTIVATES PROCOLLAGENASE.
CC - CATALYTIC ACTIVITY: Similar to stromelysin 1, but action on
CC collagen types III, IV and V is weak.
CC - COFACTOR: Binds 2 zinc ions per subunit, calcium (By similarity).
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
CC - SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X05083; CA28739.1; -
DR EMBL: M65253; AAA42202.1; -
DR PIR: B26403; KCRTS2.
DR HSSP: P08254; ISLM.
DR MEROPS: M10.006; -
DR InterPro: IPR000585; Hemopexin.
DR InterPro: IPR001818; Matrixin.
DR InterPro: IPR006026; Zn_MTPeptide.
DR InterPro: IPR006025; Zn_MTPeptide.
DR Pfam: PF00045; hemopexin; 4.
DR Pfam: PF00413; Peptidase_M10; 1.
DR Pfam: PF03933; Peptidase_M10_N; 1.

```

```

DR PRINTS; PRO0138; MATRIXIN.
DR SMART; SM00120; HX; 3.
DR SMART; SM00235; ZMC; 1.
DR PROSITE; PS00042; ZINC_PROTEASE; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
KW Collagen degradation; Extracellular matrix; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 99 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 100 476 STROMELYSIN-2.
FT DOMAIN 286 476 HEMOPEXIN-LIKE.
FT SITE 92 92 CYSTEINE_SWITCH (BY SIMILARITY).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 219 219 BY SIMILARITY.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 289 476 BY SIMILARITY.
SQ SEQUENCE 476 AA; 54221 MW; B5568F81D8BAYEE CRC64;

Query Match 8.6%; Score 187; DB 1; Length 476;
Best Local Similarity 26.4%; Pred. No. 1.7e-08;
Matches 52; Conservative 23; Mismatches 48; Indels 74; Gaps 8;

Qy 21 AATPLPPTSAHGNVAEGETKPDVTERCSQSGMSFDATTLDDNGTLPFKGEFWKSHK 80
Db 271 ATVPVPSVS-----PKETPVKCDPALSPDAVTM-LRGYLFKDKHFWRRITQ 318
Qy 81 WDRE---LISERWKVPSPVDAAFRQGH-NSVFLIKGDKVWV----- 118
Db 319 WNEPEFHLSAIFWPSYLDAAVEANNNKDRVLIKFGSQFPAVNRGENVOAGYPRKIIHTL 378
Qy 119 -YRP-----EKK-----EKGYPKLLDPEFGIPSPDLA 145
Db 379 GFPPYTKKIDAAVEEKKKTYFVVDKWRFPDTRQLMDKGRPLITDDFPIEQVDA 438
Qy 146 AVECHRGECAEGVLFP 162
Db 439 VL-----HAFGFYF 448

RESULT 14
MM01 RABIT STANDARD; PRT; 468 AA.
ID MM01 RABIT
AC P13943;
DT 01-JAN-1990 (Rel. 13, Created).
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interstitial collagenase precursor (EC 3.4.24.7) (Matrix
DE metalloproteinase-1) (MMP-1).
GN MMP1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Synovial cell;
RX MEDLINE=88077876; PubMed=2825772;
RA Fani M.E., Plucinska I.M., Mayer A.S., Gross R.H., Brinckerhoff C.E.;
RT "A gene for rabbit synovial cell collagenase: member of a family of
RT metalloproteinases that degrade the connective tissue matrix."
RL Biochemistry 26:6156-6165(1987).
RN [2]
RP SEQUENCE OF 449-468 FROM N.A.
RX STRAIN=New Zealand white;
RX MEDLINE=87029174; PubMed=3021384;
RA Fani M.E., Austin S.D., Holt P.T., Ruby P.L., Gross R.H., White H.D.,
RA Brinckerhoff C.E.;
RT "Homology between exon-containing portions of rabbit genomic clones
RT for synovial cell collagenase and human foreskin and synovial cell
RT mRNA."
RL Coll. Relat. Res. 6:239-248(1986).

```

```

FT TURN 123 124
FT HELIX 127 144
FT STRAND 148 151
FT STRAND 159 164
FT STRAND 182 184
FT TURN 190 193
FT STRAND 195 198
FT TURN 199 200
FT STRAND 203 204
FT STRAND 210 211
FT HELIX 212 224
FT TURN 225 225
FT TURN 232 233
FT TURN 235 236
FT HELIX 253 263
SQ SEQUENCE 477 AA; 54190 MW; 361CE1427E09A272 CRC64;

Query Match 8.9%; Score 195.5; DB 1; Length 477;
Best Local Similarity 27.3%; Pred. No. 3.2e-09;
Matches 67; Conservative 24; Mismatches 89; Indels 65; Gaps 11;

OY 27 PPTSAHGNVAEGTSPD-PDVTGRCSDGWSFDA-TTLDNGTMLPFKGEFVKSHKMDRE 84
DB 266 PPAESDSEVSESEPPAPGTLAMCDPALSPDAISTL--RGEILFFKORYFWR--KTRRT 321
OY 85 L-----ISERKKNPPSPVDAAFR-QGHNSVFLKGDVWVYPPKKEKGYKLIQD-BF 136
DB 322 LVPEHFPISSFWPSPSPGIDAAVEYTSRDSVTFPGKNFMAIRGNBDAQYRGHITIGF 381
OY 137 PGISPLDAVAECRGECQAEGLFPQGHGHRNGTGHGNSHTHGPEYKRCSPHLVLSALT 196
DB 382 PPTVAKKIDAI-----FDKEKOK----- 399
OY 197 SDNHGATYAFSGTHYRLDTSRDGWH-SWP--IAHQWQGSAPVADAFSWEKLYVQGT 253
DB 400 -----TTFVBDKTKWRDEKRSQMEPGYKQIADPFGLSKDLAAFESEGFYFSGS 453
OY 254 QVYVF 258
DB 454 SQFEF 458

RESULT 12
MM10 HUMAN STANDARD; PRT; 476 AA.
ID P09238;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Stromelysin-2 precursor (EC 3.4.24.22) (Matrix metalloproteinase-10)
DE (MMP-10) (Transin-2) (SL-2).
GN MMP10 OR STMW2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6833985; PubMed=2844164;
RA Muller D., Quantin B., Gesnel M.-C., Milon-Collard R., Abecassis J.,
RA Breathnach R.,
RT "The collagenase gene family in humans consists of at least four
RT members."
RL Biochem. J. 253:187-192(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RC MEDLINE=23388257; PubMed=12477932;
RA Straubeberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.O., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.T., Skalek U., Smallegange D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: CAN DEGRADE FIBROBLAST GELATIN OF TYPE I, III, IV,
CC AND V. WEAKLY COLLAGENS III, IV, AND V. ACTIVATES PROCOLLAGENASE.
CC -!- CATALYTIC ACTIVITY: Similar to stromelysin 1, but acts on
CC collagen types III, IV and V is weak.
CC -!- COFACTOR: Binds 2 zinc ions per subunit, calcium (By similarity).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
CC -!- SIMILARITY: Contains 1 hemopexin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: X07820; CAA30679.1; -.
DR EMBL: BC002591; AAH02591.1; -.
DR PIR: A28816; KCHUS2.
DR HSSP: P08254; 1SLM.
DR GenPe: M10.006; -.
DR MIM: 185260; -.
DR GO: GO:0005578; C:extracellular matrix; TAS.
DR GO: GO:0005615; C:extracellular space; TAS.
DR GO: GO:0008270; F:zinc ion binding activity; TAS.
DR InterPro: IPR000585; Hemopexin.
DR InterPro: IPR001818; Matrilin.
DR InterPro: IPR006026; NTA MTpeptidase.
DR InterPro: IPR006025; Zn MTpeptidase.
DR Pfam: PF00045; hemopexin; 4.
DR Pfam: PF00413; peptidase_M10; 1.
DR Pfam: PF03933; peptidase_M10_N; 1.
DR PRINTS: PR00138; MATRILIN.
DR SMART: SMO0120; HX; 4.
DR SMART: SMO0235; ZNMC; 1.
DR PROSITE: PS00024; HEMOPEXIN; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR PROSITE: PS00546; CYSTEINE_SWITCH; 1.
KW Hydroxylase; Metalloproteinase; Zinc; Zymogen; Calcium;
KW Collagen degradation; Extracellular matrix; Signal.
FT SIGNAL 1 17
FT PROPEP 18 98
FT CHAIN 99 476
FT DOMAIN 286 476
FT SITE 91 91
FT METAL 217 217
FT ACT SITE 218 218
FT METAL 221 221
FT METAL 227 227
FT METAL 289 476
SQ SEQUENCE 476 AA; 54151 MW; 516DCDDFE92A0D6 CRC64;

Query Match 8.7%; Score 189.5; DB 1; Length 476;
Best Local Similarity 28.4%; Pred. No. 1e-08;
Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

OY 25 PPTSAHGNVAEGTSPD-PDVTGRCSDGWSFDA-TTLDNGTMLPFKGEFVKSHKMDRE 83

```

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC
 DR EMBL: Y13185; CAJ73641.1; -.
 DR PIR: J06505; J06505.
 DR HSSP: P08254; ISLM.
 DR MEROPS: M10.006; -.
 DR MGD: MGI:97007; MMP10.
 DR InterPro: IPR001843; Fragilysin.
 DR InterPro: IPR000585; Hemopexin.
 DR InterPro: IPR001818; Matrxin.
 DR InterPro: IPR006026; Nzn_MTPeptidase.
 DR InterPro: IPR006025; Zn_MTPeptidase.
 DR Pfam: PF02051; Fragilysin_1.
 DR Pfam: PF00045; hemopexin_4.
 DR Pfam: PF00413; Peptidase_M10_1.
 DR Pfam: PF03933; Peptidase_M10_N; 1.
 DR PRINTS: PR00138; MATRILIN.
 DR SMART: SM00120; HX; 3.
 DR SMART: SM00235; ZmC; 1.
 DR PROSITE: PS00024; HEMOPEXIN; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR PROSITE: PS00546; CYSTEINE_SWITCH; 1.
 DR Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
 KW Collagen degradation; Extracellular matrix; Signal.
 FT SIGNAL 1 17
 FT PROPEP 1 99 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 100 476 STROMELYSIN-2.
 FT DOMAIN 286 476 HEMOPEXIN-LIKE.
 FT SITE 92 92 CYSTEINE SWITCH (BY SIMILARITY).
 FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT SITE 219 219 BY SIMILARITY.
 FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 289 476 BY SIMILARITY.
 SQ SEQUENCE 476 AA; 53911 MW; 2EB1CC41468F0AC6 CRC64;

Query Match 8.94; Score 195.5; DB 1; Length 476;
 Best Local Similarity 26.34; Pred. No. 3.2e-09;
 Matches 65; Conservative 26; Mismatches 85; Indels 71; Gaps 10;

QY 23 ATPLPTSAHGVNAEGTKEPDPTTERSCSDGWSFDA-TLLDNGTMLFFKGSEFVMSHKV 81
 Db 271 ATVPVLS-----VSPRPETPDKCDALSDSSTL--RGEVLFVFDKRFMRSHV 319
 QY 82 DRE---LISERKNPSPVDAAFROGH--SVFLIGDKVWVYPPKKKGYPKLQD- 134
 Db 320 NPEPEFLISAFWPTLPSDIDAAY-EAHNTDSVLIFKGSQFMVAVRGNEVQAGYPKGIHTL 378
 QY 135 EFGGIPSLDAAYECHRGEGQAEGLVLFQGHGRNNGHNSHTHGEVWRCSPHVLVLSA 194
 Db 379 GFPPYTKKIDAAV-----FEKEKK----- 398
 QY 195 LTSDNMGATYAFSGTHYRLDTSR---DGNHSWPIAHQWQSPASVADAFAFSWEKLYLVQ 251
 Db 399 -----TYFFGDKVMRFDKTRHWDKGRPRQITDDPGIIEQVDVALHGGFFFR 450
 QY 252 GTQVYVF 258
 Db 451 GSSQFEP 457

RESULT 11
 MM03_HORSE STANDARD; PRT; 477 AA.
 AC Q28397;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Stromelysin-1 precursor (EC 3.4.24.17) (Matrix metalloproteinase-3) (MMP-3).
 GN MMP3.
 OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OK NCBI_TaxID=9796;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cartilage;
 RX MEDLINE=99074117; PubMed=9858406;
 RA Richardson D.W., Dodge G.R.;
 RT "Molecular characteristics of equine stromelysin and the tissue
 RT inhibitor of metalloproteinase 1.";
 RL Am. J. Vet. Res. 59:1557-1562(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cartilage;
 RX MEDLINE=98104320; PubMed=9442239;
 RA Ballman C.E., Nixon A.J.;
 RT "Molecular cloning and cartilage gene expression of equine stromelysin
 RT 1 (matrix metalloproteinase 3).";
 RL Am. J. Vet. Res. 59:30-36(1998).
 RN (3)
 RP 3D-STRUCTURE MODELING.
 RA Mallena S.C., Sharma J.A.R.P.;
 RT "Theoretical model of horse stromelysin.";
 RL Submitted (MAR-2002) to the PDB data bank.
 CC -1- FUNCTION: CAN DEGRADE FIBRONECTIN, LAMININ, GELATINS OF TYPE I,
 CC III, IV, AND V; COLLAGENS III, IV, X, AND IX, AND CARTILAGE
 CC PROTEOLYCN'S. ACTIVATES PROCOLLAGENSE.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage where P1', P2' and P3'
 CC are hydrophobic residues.
 CC -1- COPACITOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
 CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL: U62529; AAB05774.1; -.
 DR PDB: 1L91; 03-APR-02.
 DR MEROPS: M10.005; -.
 DR InterPro: IPR000585; Hemopexin.
 DR InterPro: IPR001818; Matrxin.
 DR InterPro: IPR006026; Nzn_MTPeptidase.
 DR InterPro: IPR006025; Zn_MTPeptidase.
 DR Pfam: PF00045; hemopexin_4.
 DR Pfam: PF00043; Peptidase_M10_1.
 DR Pfam: PF03933; Peptidase_M10_N; 1.
 DR PRINTS: PR00138; MATRILIN.
 DR SMART: SM00120; HX; 4.
 DR SMART: SM00235; ZmC; 1.
 DR PROSITE: PS00024; HEMOPEXIN; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR PROSITE: PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 KW Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 KW Collagen degradation; Extracellular matrix; Signal; 3D-structure.
 FT SIGNAL 1 17
 FT PROPEP 1 99 ACTIVATION PEPTIDE.
 FT CHAIN 100 477 STROMELYSIN-1.
 FT DOMAIN 287 477 HEMOPEXIN-LIKE.
 FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT SITE 219 219 BY SIMILARITY.
 FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DISULFID 290 477 BY SIMILARITY.
 FT CONFLICT 346 346 V -> E (IN REF. 2).
 FT TURN 104 105
 FT STRAND 113 118

```

Db 337 T--HDAVA-QIRGAFFFKGYFRLTR-DRHLVSLDPAQWHRFWRGLPLHLSVDVAV 392
OY 102 -ROGNSVFLKGDVKVWYPPREKKEGYKLLDPEPGIPSLDAVAECHGECAEVL 160
Db 393 ERTSDHKVFRKGRDYMFKDNNVEGYPRVSD-FSLPPGIDDA----- 437
OY 161 FFOGHNGTGHGNSTHHGPEYMRCSPHLVLSALTSDNHGATYAFSGTHYRLDT-SRD 219
Db 438 FSWAHNDR-----TYFFKQQLVWRDYDHTRH 463
OY 220 GWHSMPT-AHQWPGPSAVDAAFSWE-KLYVGSTQYVYPLTKGTYLVSGYPRLEKE 277
Db 464 MDGPYPAGSPLMRGVPSLTLDAMRWSDGASYFRFG-QEYWKVLDELEVAQVPGPOSTARD 522
OY 278 -----VGFPHGIILSDVDA 292
Db 523 WLVCDSQADSGSVAAGVDA 542

RESULT 9
MM12_RABIT
ID_MM12_RABIT STANDARD; PRT; 464 AA.
AC P79227;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Macrophage metalloelastase precursor (EC 3.4.24.65) (MME) (Matrix
  metalloproteinase-12) (MMP-12).
GN MMP12.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand;
RA Hou P.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
  SIGNIFICANT ELASTOLYTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of soluble and insoluble elastin.
  Specific cleavages are also produced at 14-Ala-[Leu-15 and 16-
  Tyr-[Leu-17 in the B chain of insulin.
CC -1- COFACTOR: Binds 2 zinc ions per subunit, calcium (By similarity).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U88652; AAB46993.1; -
DR HSSP; P03956; JGCL.
DR MEROPS; M10.009; -.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Matrxin.
DR InterPro; IPR006026; N2A_MTPeptide.
DR InterPro; IPR006025; Zn_MTPeptide.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; peptidase_M10; 1.
DR Pfam; PF03933; peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; Zmnc; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
KW Hydroxylase; Metalloproteinase; Glycoprotein; Zinc; Zymogen; Calcium;

```

```

KW Extracellular matrix; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPE 18 100 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 101 464 MACROPHAGE METALLOELASTASE.
FT DOMAIN 274 464 HEMOPEXIN-LIKE.
FT SITE 87 87 CYSTEINE SWITCH (BY SIMILARITY).
FT METAL 213 213 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 217 217 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 223 223 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 69 69 N-LINKED (GLCNAC... ) (POTENTIAL).
FT DISULFID 277 464 BY SIMILARITY.
SQ SEQUENCE 464 AA; 53584 MW; BC8398FD97A15E CRC64;

Query Match 9.0%; Score 196; DB 1; Length 464;
Best Local Similarity 27.4%; Pred. No. 2.8e-09;
Matches 64; Conservative 19; Mismatches 91; Indels 60; Gaps 8;

OY 30 SAHGNVABGETKPPDYTE--RCSDGSPDATYLLDNGTMLFPKGEFW-----KSHKMDR 83
Db 255 SLVGGPEHQHMPKPDNPEPTACDHLKFDVTTVGN-KIFFFQDSFFWMKIPSSSTTSV 313
OY 84 ELISERKKNPPSPYDAAFROG-HNSVFLKGDVKVWYPPREKKEGYKLLD-EFGIIPS 141
Db 314 RLISLWPTLPSGIEAAVEIGDRHQVLPFKGDKFWLISHRLQPNYFKSISHSGFPFVK 373
OY 142 PLDAVAECHGECAEVLFFQGHGHNHGCHGNSTHHGPEYMRCSPHLVLSALTSNNG 201
Db 374 KIDNAV-----FNSPL-----R 385
OY 202 ATYAFSGTHYRLDTSR---DGMHSMPIAHQWPGPSAVDAAFSWEKLYLVQG 252
Db 386 KTYFVNVNLRVYRERENVADAGYPKLITKHPEIGKIDAVFQFYVYVFGQ 439

RESULT 10
MM10_MOUSE
ID_MM10_MOUSE STANDARD; PRT; 476 AA.
AC O55123;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Stromelysin-2 precursor (EC 3.4.24.22) (Matrix metalloproteinase-10)
  (MMP-10) (Transin-2) (SL-2).
GN MMP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA MEDLINE=98087420; Pubmed=9427548;
RX Madlener M., Werner S.;
RT "cDNA cloning and expression of the gene encoding murine stromelysin-2
  (MMP-10).";
RT Gene 202:75-81(1997).
CC -1- FUNCTION: CAN DEGRADE FIBRONECTIN, GELATINS OF TYPE I, III, IV,
  AND V, WEAKLY COLLAGENS III, IV, AND V. ACTIVATES PROCOLLAGENASE.
CC -1- CATALYTIC ACTIVITY: Similar to stromelysin 1, but action on
  collagen types III, IV and V is weak.
CC -1- COFACTOR: Binds 2 zinc ions per subunit, calcium (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SMALL INTESTINE. WEAK LEVELS IN
  HEART AND LUNG.
CC -1- INDUCTION: By wounding.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial

```

RESULT 8
MM17_HUMAN STANDARD; PRT; 606 AA.
ID 09UL29; Q14850;
AC 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Matrix metalloproteinase-17 precursor (EC 3.4.24.-) (MMP-17)
DE (Membrane-type matrix metalloproteinase 4) (MT-MMP 4) (Membrane-type-4
DE matrix metalloproteinase) (MT4-MMP).
GN MMP17 OR MT4MMP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Monocytic leukemia;
RX MEDLINE=99402951; PubMed=10471807;
RA Kajita M., Kinoh H., Ito N., Takamura A., Itoh Y., Okada A., Sato H.,
RA Seiki M.;
RT "Human membrane type-4 matrix metalloproteinase (MT4-MMP) is encoded
RT by a novel major transcript: isolation of complementary DNA clones
RT for human and mouse mt4-mmp transcripts.";
RL FEBS Lett. 457:353-356(1999).
RP REVISIONS TO 44; 47; 205-210; 224 AND 228.
RA Seiki M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Breast carcinoma;
RX MEDLINE=96234364; PubMed=8640782;
RA Puente X.S., Pendas A.M., Llano E., Velasco G., Lopez-Otin C.;
RT "Molecular cloning of a novel membrane-type matrix metalloproteinase
RT from a human breast carcinoma.";
RL Cancer Res. 56:944-949(1996).
RN [4]
RP SEQUENCE OF 129-302 FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20020281; PubMed=10551873;
RA Wang Y., Johnson A.R., Ye Q.-Z., Dyer R.D.;
RT "Catalytic activities and substrate specificity of the human membrane
RT type 4 matrix metalloproteinase catalytic domain.";
RL J. Biol. Chem. 274:33043-33049(1999).
RN [5]
RP GPI-ANCHOR.
RX MEDLINE=20036570; PubMed=10567400;
RA Itoh Y., Kajita M., Kinoh H., Mori H., Okada A., Seiki M.;
RT "Membrane type 4 matrix metalloproteinase (MT4-MMP, MMP-17) is a
RT glycosylphosphatidylinositol-anchored proteinase.";
RL J. Biol. Chem. 274:34260-34266(1999).
RN [6]
RP CHARACTERIZATION.
RX MEDLINE=20008793; PubMed=10543448;
RA Koltenbrock H., Essere L., Ubrich N., Will H.;
RT "Biochemical characterization of the catalytic domain of membrane-type
RT 4 matrix metalloproteinase.";
RL Biol. Chem. 380:1103-1108(1999).
CC -1- FUNCTION: Endopeptidase that degrades various components of the
CC extracellular matrix, such as fibrin. May be involved in the
CC activation of membrane-bound precursors of growth factors or
CC inflammatory mediators, such as tumor necrosis factor-alpha. May
CC also be involved in tumoral process. Not obvious if able to
CC proteolytically activate progelatinase A. Does not hydrolyze
CC collagen types I, II, III, IV and V, gelatin, fibronectin,
CC laminin, decorin nor alpha1-antitrypsin.
CC -1- CATALYTIC ACTIVITY: Cleaves pro-TNF-alpha at the 74-Ala-Gln-75
CC site.
CC -1- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;

CC Name=Long;
CC IsoId=09UL29-1; Sequence=Displayed;
CC Name=Short; Synonyms=Puente;
CC IsoId=09UL29-2; Sequence=VSP 005456;
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LEUKOCYTES, COLON, OVARY
CC TESTIS AND BREAST CANCER. EXPRESSED ALSO IN MANY TRANSFORMED AND
CC NON-TRANSFORMED CELL TYPES.
CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FIRIN ENDOPEPTIDASE (BY
CC SIMILARITY). BELONGS TO PEPTIDASE FAMILY M10A.
CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AB021225; BA83707.2; -.
CC EMBL: X89576; CA61753.1; -.
CC HSSP: Q02853; IHW5.
CC MEROPS: M10.017; -.
CC Genew: HGNC:7163; MMP17.
CC MIM: 602285; -.
CC GO: GO:0005887; C: integral to plasma membrane; TAS.
CC GO: GO:0008047; E: enzyme activator activity; TAS.
CC GO: GO:0004222; F: metalloendopeptidase activity; TAS.
CC GO: GO:0008270; F: zinc ion binding activity; TAS.
CC InterPro: IPR000585; Hemopexin.
CC InterPro: IPR001818; Matrxin.
CC InterPro: IPR006026; Nzn_MTPeptidse.
CC InterPro: IPR006025; Zn_MTPeptidse.
CC Pfam: PF00045; hemopexin; 4.
CC Pfam: PF00413; Peptidase_M10; 1.
CC Pfam: PF03933; Peptidase_M10_N; 1.
CC PRINTS: PR00138; MATRXIN.
CC SMART: SM00235; ZMNC; 1.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC PROSITE: PS00546; CYSTEINE_SWITCH; FALSE NEG.
CC PROSITE: PS00024; HEMOPEXIN; FALSE NEG.
CC KX Hydroxylase; Metalloproteinase; Zinc; Calcium; Zymogen; Signal;
CC KX GPI-anchor; Extracellular matrix; Alternative splicing.
CC FT SIGNAL 1 38
CC FT PROPEP 39 128 BY SIMILARITY.
CC FT CHAIN 129 568 MATRIX METALLOPROTEINASE-17.
CC FT PROPEP 569 606 REMOVED IN MATURE FORM (POTENTIAL).
CC FT DOMAIN 332 529 HEMOPEXIN-LIKE.
CC FT SITE 113 113 CYSTEINE SWITCH (POTENTIAL).
CC FT METAL 251 251 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT ACT_SITE 252 252 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 255 255 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 261 261 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT DISULFID 335 526 BY SIMILARITY.
CC FT LIPID 568 568 GPI-ANCHOR (POTENTIAL).
CC FT CARBOHYD 125 128 POLY-ARG.
CC FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARSPLIC 1 87 Missing (in isoform Short).
CC FT VARSPLIC 1 87 /FTId=VSP 005456.
CC SQ SEQUENCE 606 AA; 67006 MW; P5D51445118F4DF5 CRC64;
Query Match 9.1%; Score 199; DB 1; Length 606;
Best Local Similarity 25.0%; Pred. No. 2, 1e-09;
Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;
QY 5 LGAPVALG-----WELCMSLAATPL-----PPTSAHGNGVAGETKPPDPYERGS 51
DB 277 VGPPLRYGGLPYEDKRVWOLYGVRESVSPPTAPBEPPLLPDPDSSAPPRKDVHRS 336
QY 52 DGWSPDATTLDNGTALPFKSGFVWVKSHKMDRELISER-----WKNFP---SPVDAAF 101

```

QY 134 DEFGIPSPALDAVECHRGCEQAEGLFPQGHGHNCTGHNSTHNGEYWCSPHLYLS 193
DB 427 D-FSLPFGGIDA-----VFMAHND-----446
QY 194 ALTSNHNATYAFSGTHYWRD--TSR--DGWHS-WPIAHQWPGQPSAVDAAFSWEK-KL 247
DB 447 -----TYFFKQQLWRYDDHTRRDPGYPAGGPI--WRGVPMLDAMRWSDGAS 494
QY 248 YLVQGVYVFLTKGTYLVSGYPRLEKE--VGTPHGILLDSVDAAFICGSSRLHIM 304
DB 495 YFFRG-QGYWKVLDGLAEARPGYPOSTRDMVLCSE-----LADADVGGPO-----542
QY 305 AGRLMLDLKSGAGATWTELPWPEKVDG-ALCMKSLGPNSSCANGPGLYLHGPMLY 363
DB 543 -GR-----SGAQ-----DGLAVCS-----555
QY 364 CYSDVEKNAKALPQPNVTSL 387
DB 556 CTSDAHRL-----ALPSLLLTPL 575

RESULT 7
MM12_HUMAN
ID MM12_HUMAN STANDARD; PRT; 470 AA.
AC P39900;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Macrophage metalloelastase precursor (EC 3.4.24.65) (HME) (Matrix
DE metalloproteinase-12) (MMP-12) (Macrophage elastase) (ME).
GN MMP12 OR HME.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N. A.
RC TISSUE=Alveolar macrophage;
RX MEDLINE=94043200; PubMed=8226919;
RA Shapiro S.D., Kobayashi D.K., Ley T.J.;
RT "Cloning and characterization of a unique elastolytic
RT metalloproteinase produced by human alveolar macrophages.";
RL J. Biol. Chem. 268:23824-23829 (1993).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=9727370; PubMed=9115292;
RA Gronski T.J., Jr., Martin R.L., Kobayashi D.K., Walsh B.C.,
RA Holman M.C., Huber M., Van Wart H.G., Shapiro S.D.;
RT "Hydrolysis of a broad spectrum of extracellular matrix proteins by
RT human macrophage elastase.";
RL J. Biol. Chem. 272:12189-12194 (1997).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 100-280.
RX MEDLINE=21460859; PubMed=11575929;
RA Nar H., Werle K., Bauer M.M.T., Dollinger H., Jung B.;
RT "Crystal structure of human macrophage elastase (MMP-12) in complex
RT with a hydroxamic acid inhibitor.";
RL J. Mol. Biol. 312:743-751 (2001).
CC -1- FUNCTION: May be involved in tissue injury and remodeling. Has
CC significant elastolytic activity. Can accept large and small amino
CC acids at the P1' site, but has a preference for leucine. Aromatic
CC or hydrophobic residues are preferred at the P1 site, with small
CC hydrophobic residues (preferably alanine) occupying P3.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of soluble and insoluble elastin.
CC Specific cleavages are also produced at 14-Ala-[Leu-15 and 16-
CC Tyr-1-Leu-17 in the B chain of insulin.
CC -1- COFACTOR: Requires calcium and zinc for activity.
CC -1- TISSUE SPECIFICITY: FOUND IN ALVEOLAR MACROPHAGES BUT NOT IN
CC PERIPHERAL BLOOD MONOCYTES.
CC -1- INDUCTION: BY EXPOSURE TO LYPOPOLYSACCHARIDE. INHIBITED BY
CC DEXAMETHASONE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.

```

```

CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
CC -----
DB EMBL, L23808; AA58658.1; ALT_SEQ.
DB PIR, A49499; A49499.
DB PDB, 1JIZ; 03-JUL-02.
DB PDB, 1JK3; 28-SEP-01.
DB MEROPS, M10.009; -.
DB GeneW, HGNC:7158; MMP12.
DB MIM, 601046; -.
DB GO, GO:0004234; P:macrophage elastase activity; TAS.
DB GO, GO:0008270; F:zinc ion binding activity; TAS.
DB GO, GO:0006508; P:proteolysis and peptidolysis; TAS.
DB InterPro, IPR00585; Hemopexin.
DB InterPro, IPR001818; Matrixin.
DB InterPro, IPR006026; NZn_MTPeptide.
DB InterPro, IPR005025; Zn_MTPeptide.
DB Pfam, PF00045; hemopexin; 4.
DB Pfam, PF00933; peptidase M10_N; 1.
DB PRINTS, PR00138; MATRIXIN.
DB SMART, SM00120; HX; 4.
DB SMART, SM00235; ZnMC; 1.
DB PROSITE, PS00024; HEMOPEXIN; 1.
DB PROSITE, PS00142; ZINC_PROTEASE; 1.
DB PROSITE, PS00546; CYSTEINE_SWITCH; 1.
DB HydroLase, Metalloproteinase, Glycoprotein; Zinc; Zymogen; Calcium;
KW Extracellular matrix; Signal; 3D-structure.
FT SIGNAL 1 16
FT PROPEP 17 105
FT CHAIN 106 470
FT DOMAIN 279 470
FT SITE 92 92
FT METAL 218 218
FT ACT_SITE 219 219
FT METAL 222 222
FT METAL 228 228
FT CARBOHYD 20 20
FT CARBOHYD 285 285
FT DISULFID 282 470
SQ SEQUENCE 470 AA; 54001 MW; 8C745E8A8CDEA216 CRC64;

Query Match 9.3%; Score 202.5; DB 1; Length 470;
Best Local Similarity 27.3%; Pred. No. 7, 6e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGETKDDPYTE--RCSDGWSFDATLIDNGTMLPKGEFVW-----KSHKWR 83
DB 260 SLVDDPKENQRLPNDNSEBALCPDNLSDAVTVGN-KLFFXDRFWMKVSERPTSV 318
QY 84 ELISERKKNPSPVDAFR-QGHSVFLIKGDKWVYPRPEKKEGYKLLQD-EFPGIPS 141
DB 319 NLISLWPLTPSGIEAAYEIEARQVFLPFDQKWLISNLRPEVNYPKSIHSFGFPVFK 378
QY 142 PLDAVCHRGCEQAEGLFPQGHGHNCTGHNSTHNGEYWCSPHLYLSALTSNMHG 201
DB 379 KIDAAV-----FNRRFR-----391
QY 202 ATYAFSGTHYWRDTSRD---GWSWPIAHQWPGQPSAVDAAFSWEK-LYLVQGVY 256
DB 392 -TYFFVQNWRYRDERQWMDPGYPKLITGNPGSIGPK-IDAVFYSKNKYYFFQGSNGF 449
QY 257 VF 258
DB 450 EY 451

```


QY 119 YPEEKEKGYPKLLQDEFPPIPLDPAVECHRECOAGVLFPGQ----- 164
 DB 121 YPRE-KEKENPRELQGEFPVPSPLDAAVCHRGECODESVLFPGQHTWTFMDSTTKTK 179
 QY 165 -----HGH-----RNGTGH 173
 DB 180 ERLMPAVNGCSSAMRWISRYCYCRGNQFLRFPDVTGHVDPKYPDRVDYFMSCPRGHAH 239
 QY 174 GNSHTHGPXYMRCSPHLYSALTSDNHGATYASGTHYMLDPSRDQMSWPIAHQWPOQ 233
 DB 240 RAAHTHGDD--RCSPLVLTALLSDNHGATYARGHYHRLDTSRDQMSWPIAHQMSHG 297
 QY 234 PSAVDAFAFSEWEEKLYVQGTQVYVFLTKGQYTLVSGYPKLEKXVGTPHGILLSDVAAF 293
 DB 298 PSAVDAFAFSEWEEKLYVQGTQVYVFLTKGQYTLVSGYPKLEKXVGTPHGILLSDVAAF 357
 QY 294 ICGSSRLHIMAGRLMLWDLKSGAQTWTELFWPHKXGALCMKSLGPNCSANGRC 353
 DB 358 VCPGSTRHVMARCKLWMDLSLGAQGPWTELPWPHKXGALCMKSLGPNCSANGRC 417
 QY 354 LVLIHGPNYCYSDVEKLNAAKALPOPOVNTSLGCG 389
 DB 418 LVHGPNYCYSDVEKLNAAKALPOPOVNTSLGCG 453

RESULT 6
 MM17_MOUSE STANDARD; PRT; 578 AA.
 AC 09R053;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Matrix metalloproteinase-17 precursor (EC 3.4.24.-) (MMP-17)
 DE (Membrane-type matrix metalloproteinase 4) (MT-MMP 4) (Membrane-type-4
 matrix metalloproteinase) (MT4-MMP).
 GN MMP17 OR MT4MP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic brain;
 RX MEDLINE=99402951; PubMed=10471807;
 RA Kajita M., Kinoh H., Ito N., Takamura A., Itoh Y., Okada A., Sato H.,
 RT "Human membrane type-4 matrix metalloproteinase (MT4-MMP) is encoded
 by a novel major transcript: isolation of complementary DNA clones for
 human and mouse mt4-mmp transcripts.";
 RL FEBS Lett. 457:353-356(1999).
 RN [2]
 RP REVISIONS TO C-TERMINUS.
 RA Seiki M.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC TISSUE=Brain;
 RX MEDLINE=20261520; PubMed=10799478;
 RA English W.R., Puente X.S., Freije J.M.P., Knaeuper V., Amour A.,
 RA Merryweather A., Lopez-Otin C., Murphy G.;
 RT "Membrane type 4 matrix metalloproteinase (MMP17) has tumor necrosis
 factor-alpha convertase activity but does not activate pro-MMP2.";
 RL J. Biol. Chem. 275:14046-14055(2000).
 RN [4]
 RP GPI-ANCHOR, AND MUTAGENESIS OF GLU-248.
 RX MEDLINE=20036570; PubMed=10567400;
 RA Itoh Y., Kajita M., Kinoh H., Mori H., Okada A., Seiki M.;
 RT "Membrane type 4 matrix metalloproteinase (MT4-MMP, MMP-17) is a
 glycosylphosphatidylinositol-anchored proteinase.";
 RL J. Biol. Chem. 274:34260-34266(1999).
 CC - FUNCTION: Endopeptidase that degrades various components of the
 CC extracellular matrix, such as fibrin. May be involved in the
 CC activation of membrane-bound precursors of growth factors or

CC inflammatory mediators, such as tumor necrosis factor-alpha. May
 CC also be involved in tumoral processes. Not obvious if able to
 CC proteolytically activate progelatinase A. Does not hydrolyze
 CC collagen types I, II, III, IV and V, gelatin, fibronectin,
 CC laminin, decorin nor alpha1-antitrypsin.
 CC - CATALYTIC ACTIVITY: Cleaves pro-TNF-alpha at the 74-Ala-I-Gln-75
 CC site.
 CC - COFACTOR: Binds 1 zinc ion per subunit, calcium (by similarity).
 CC - SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC - TISSUE SPECIFICITY: EXPRESSED BY MONOCYTES AND MACROPHAGES.
 CC - PTM: THE PRECURSOR IS CLEAVED BY A PURIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
 CC - SIMILARITY: Contains 1 hemopexin-like domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL; AB021224; BAA82708.2; -;
 DR EMBL; AJ010731; CAB92315.1; -;
 DR HSSP; Q02853; 1HV5.
 DR MEROPS; M10.017; -;
 DR MGD; MG1:146076; MMP17.
 DR InterPro; IPR000385; Hemopexin.
 DR InterPro; IPR001818; Matrixin.
 DR InterPro; IPR006026; Nzn_MTPeptide.
 DR InterPro; IPR006025; Zn_MTPeptide.
 DR Pfam; PF00045; hemopexin; 4.
 DR Pfam; PF00413; peptidase_M10; 1.
 DR Pfam; PF03933; peptidase_M10_N; 1.
 DR PRINTS; PR00136; MATRIXIN.
 DR SMART; SM00120; HX; 4.
 DR SMART; SM00235; Zmnc; 1.
 DR PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
 DR PROSITE; PS00024; HEMOPEXIN; FALSE NEG.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloproteinase; Zinc; Calcium; Zymogen; Signal;
 KW GPI-anchor; Extracellular matrix.
 KM SIGNAL 1 39
 FT PROPEP 125 558
 FT CHAIN 125 558
 FT PROPEP 559 578
 FT DOMAIN 333 530
 FT SITE 109 109
 FT METAL 247 247
 FT ACT SITE 248 248
 FT METAL 251 251
 FT METAL 251 251
 FT METAL 257 257
 FT DISULFID 336 527
 FT LIPID 558 558
 FT CARBOHYD 136 136
 FT CARBOHYD 322 322
 FT MUTAGEN 248 248
 FT CONFLICT 45 45
 FT CONFLICT 277 277
 FT SEQUENCE 578 AA; 64319 MW; 09AA933B63A02178 CRC64;
 SO

Query Match 9.3%; Score 203; DB 1; Length 578;
 Best Local Similarity 25.0%; Pred. No. 9, 1e-10;
 Matches 96; Conservative 33; Mismatches 109; Indels 146; Gaps 24;

QY 25 PLPTPSAHGNVAGETKPPDYTERSCDSWSPDATTLDNGTMLFPKGEFWKSHKMBE 84
 DB 317 PEPPNN-----RSSRPCKDVPHRCTA--HFDVAVA-QRGEAFEFKGYFWRLTR-DRH 366
 QY 85 LISR-----WKQPF--SPVDAF-RQGNVSVLIGDKQWVYVPEKKEKGYPKLQ 133
 DB 367 LVSLQPAQMRHFWRGDPLHLSDVDAVETSDHKIVFFKGRDWMFKDNNVEGYPKPV 426

KM Glycoprotein; Heme; Plasma; Repeat; Transport; Signal.
 FT SIGNAL 1 23
 FT CHAIN 1 23
 FT DOMAIN 24 460
 FT DOMAIN 56 93
 FT DOMAIN 97 140
 FT DOMAIN 187 230
 FT DOMAIN 261 304
 FT DOMAIN 306 349
 FT METAL 79 79
 FT METAL 149 149
 FT DISULFID 50 230
 FT DISULFID 148 153
 FT DISULFID 187 199
 FT DISULFID 255 458
 FT DISULFID 364 406
 FT CARBOHYD 416 433
 FT CARBOHYD 38 38
 FT CARBOHYD 64 64
 FT CARBOHYD 186 186
 FT CARBOHYD 240 240
 FT CARBOHYD 246 246
 FT CARBOHYD 419 419
 FT CONFLICT 30 30
 FT CONFLICT 233 233
 FT CONFLICT 453 453
 SQ SEQUENCE 460 AA; 51340 MW; 363BAC7520D1E39 CRC64;

Query Match 70.7%; Score 1545.5; DB 1; Length 460;
 Best Local Similarity 62.9%; Pred. No. 5e-125;
 Matches 292; Conservative 38; Mismatches 55; Indels 79; Gaps 6;

QY 1 MARVLGAPVALGLMSLWLSLAIAATPLPTSAHGNVAGEB--TKDDPVTETKSCDGSWFD 58
 DB 1 MARVLGAPVALGLMSLWLSLAIAATPLPTSAHGNVAGEB--TKDDPVTETKSCDGSWFD 58
 QY 59 TITLDNGTMLFFKGEFVSKHMDRELISERWKPPSPVDAAFROGANSVFLIKGDKWV 118
 DB 59 ATMDHNGTMLFFKGEFVSKHMDRELISERWKPPSPVDAAFROGANSVFLIKGDKWV 117
 QY 119 YPEKKEKGYKPLQDEFFPGIPSPDLAAVECHRGCEQAGVLPFGQH----- 165
 DB 118 YPEKKEKGYKPLQDEFFPGIPSPDLAAVECHRGCEQAGVLPFGQH----- 177
 QY 166 -----GH-- 167
 DB 166 -----GH-- 167
 QY 178 ERSWSTVGNCTALRWLERYCFOGNKFLRFPVTVGEVPRYPLDARDYFVSCRGHGR 237
 DB 178 ERSWSTVGNCTALRWLERYCFOGNKFLRFPVTVGEVPRYPLDARDYFVSCRGHGR 237
 QY 168 -RNGTGNGSTHGHGPEYKRCSPHLVLSALJSDNHGATYAEFGTHWRIDTSRDGHSWPI 226
 DB 238 PRNGTNGSTHGHGPEYKRCSPHLVLSALJSDNHGATYAEFGTHWRIDTSRDGHSWPI 295
 QY 227 AHWPQGPSTVDAAFSMWEKLYLVQGTQVYVFLTKGGTIVLSGPKRLKEKVGTPHGIIL 286
 DB 296 AHWPQGPSTVDAAFSMWEKLYLVQGTQVYVFLTKGGTIVLSGPKRLKEKVGTPHGIIL 355
 QY 287 DSVDAFICPSSSLHIMAGRLWMLDLKSGAQTWELPWPHEKVGALCMESLIGNS 346
 DB 356 ETDAFASCPSSSLYVSSGRRLWMLDLKSGAQTWELPWPHEKVGALCMESLIGNS 415
 QY 347 CSANGPGLYLHNGNLVYSDVEKLNAAKALPQONTSLIGCT 390
 DB 416 CSANGPGLYLHNGNLVYSDVEKLNAAKALPQONTSLIGCT 459

RESULT 5
 HEMO_PIG STANDARD; PRT; 459 AA.
 AC P50828;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 GN Hemopexin precursor (Hyaluronidase) (EC 3.2.1.35).
 DN HPX.

OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-48 AND 368-388.
 RC TISSUE=Liver;
 RX MEDLINE=95096047; PubMed=7798203;
 RA Zhu L., Hope T.J., Hall J., Davies A., Stern M.,
 RA Mueller-Eberhard U., Stern R., Parslow T.G.;
 RT "Molecular cloning of a mammalian hyaluronidase reveals identity with
 RT hemopexin, a serum heme-binding protein."
 RL J. Biol. Chem. 269:32092-32097(1994).
 CC -1- FUNCTION: Binds heme and transports it to the liver for breakdown
 CC and iron recovery, after which the free hemopexin returns to the
 CC circulation.
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1->4-linkages between N-
 CC acetyl-beta-D-glucosamine and D-glucuronate residues in
 CC hyaluronate.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
 CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U14751, AAC48457.1; -.
 CC PIR: A55486; A55486.
 CC HSSP: P20058; IHN.
 CC InterPro: IPR00585; Hemopexin.
 CC Pfam: PF00045; hemopexin; 5.
 CC SMART: SM00120; HX; 5.
 CC PROSITE: PS00024; HEMOPEXIN; 2.
 KW Glycoprotein; Heme; Plasma; Repeat; Transport; Signal; Hydrolase;

KM SIGNAL 1 28
 FT CHAIN 29 459
 FT DOMAIN 58 95
 FT DOMAIN 99 142
 FT DOMAIN 189 232
 FT DOMAIN 256 299
 FT DOMAIN 301 344
 FT METAL 81 81
 FT METAL 151 151
 FT DISULFID 52 232
 FT DISULFID 150 155
 FT DISULFID 189 201
 FT DISULFID 250 453
 FT DISULFID 359 401
 FT CARBOHYD 411 428
 FT CARBOHYD 40 40
 FT CARBOHYD 82 82
 FT CARBOHYD 188 188
 FT CARBOHYD 241 241
 SQ SEQUENCE 459 AA; 51305 MW; DB06B44C29789CF CRC64;

Query Match 70.7%; Score 1544; DB 1; Length 459;
 Best Local Similarity 63.4%; Pred. No. 6.8e-125;
 Matches 289; Conservative 33; Mismatches 64; Indels 70; Gaps 5;

QY 1 MARVLGAPVALGLMSLWLSLAIAATPLPTSA--HGNVAGEBTKDDPVTETKSCDGSWFD 58
 DB 1 MARVLGAPVALGLMSLWLSLAIAATPLPTSA--HGNVAGEBTKDDPVTETKSCDGSWFD 60
 QY 59 TITLDNGTMLFFKGEFVSKHMDRELISERWKPPSPVDAAFROGANSVFLIKGDKWV 118
 DB 61 TITLDNGTMLFFKGEFVSKHMDRELISERWKPPSPVDAAFROGANSVFLIKGDKWV 120

```

CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62642; AAA41337.1; -
DR EMBL; X60006; CAA42621.1; -
DR PIR; A43079; OORT.
DR HSSP; P20058; 1HXN.
DR InterPro; IPR000585; Hemopexin.
DR Pfam; PF00045; hemopexin; 5.
DR PROSITE; PS00024; HEMOPEXIN; 1.
KM Glycoprotein; Heme; Plasma; Repeat; Transport; Signal.
FT SIGNAL 1 23
FT CHAIN 24 460
FT DOMAIN 56 93 HEMOPEXIN-LIKE 1.
FT DOMAIN 97 140 HEMOPEXIN-LIKE 2.
FT DOMAIN 187 230 HEMOPEXIN-LIKE 3.
FT DOMAIN 261 304 HEMOPEXIN-LIKE 4.
FT DOMAIN 306 349 HEMOPEXIN-LIKE 5.
FT METAL 79 79 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 149 149 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT DISULFID 50 230 BY SIMILARITY.
FT DISULFID 148 153 BY SIMILARITY.
FT DISULFID 187 199 BY SIMILARITY.
FT DISULFID 255 458 BY SIMILARITY.
FT DISULFID 364 406 BY SIMILARITY.
FT DISULFID 416 433 BY SIMILARITY.
FT CARBOHYD 38 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 38 38 N -> C (IN REF. 2).
FT CONFLICT 49 50 HC -> KW (IN REF. 2).
SQ SEQUENCE 460 AA; 51291 MW; A1D67D05B5BFD83 CRC64;

Query Match 71.2%; Score 1556.5; DB 1; Length 460;
Best Local Similarity 63.8%; Pred. No. 5.8e-126;
Matches 296; Conservative 35; Mismatches 54; Indels 79; Gaps 6;

```

```

DB 356 DTIDAFSCPGSSKLYVNSGRRLMWLDKSGAQTAEIWSWPHKXVDCALCEKSLGPGS 415
QY 347 CGANRGVLYLHGNPLVYCSVDPEKNAKALPOPNMTSLDCT 390
DB 416 CSSNGPNLFFHGNPLVYCSVDPEKNAKALPOPNMTSLDCT 459
RESULT 4
ID HEMO_MOUSE STANDARD; PRT; 460 AA.
AC Q91X72; P97824; O8WU0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hemopexin precursor.
GN HPX OR HPXN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1] SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan A., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Maras A., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Bitterfield Y.S.N., Krzywinski M.I., Skalska U., Smilow D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2] SEQUENCE OF 7-460 FROM N.A.
RP TISSUE=Liver;
RC Koepsel R.R., Rohrbach D.H., Breckheiser B.B.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS HEME AND TRANSPORTS IT TO THE LIVER FOR BREAKDOWN
CC AND IRON RECOVERY, AFTER WHICH THE FREE HEMOPEXIN RETURNS TO THE
CC CIRCULATION.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC01246; AAH1246.1; -
DR EMBL; BC01901; AAH1901.1; -
DR EMBL; U89889; AAB49490.1; -
DR MGD; MGI:105112; Hpxn.
DR InterPro; IPR000585; Hemopexin.
DR Pfam; PF00045; hemopexin; 5.
DR SMART; SM00120; HX; 5.
DR PROSITE; PS00024; HEMOPEXIN; 1.

```

CC or send an email to license@isb-sib.ch).

 CC EMBL; X16429; CA34452.1; -.
 DR PDB; 1HXN; 15-OCT-95.
 DR PDB; 1QJ5; 03-FEB-00.
 DR InterPro; IPR000585; Hemopexin.
 DR pfam; PF00045; hemopexin; 5.
 DR SMART; SM00120; HX; 5.
 DR PROSITE; PS00024; Hemopexin; 1.
 KM Glycoprotein; Heme; Plasma; Repeat; Transport; Signal; 3D-structure.
 FT SIGNAL 1 25
 FT CHAIN 26 460
 FT DOMAIN 58 95 HEMOPEXIN-LIKE 1.
 FT DOMAIN 99 143 HEMOPEXIN-LIKE 2.
 FT DOMAIN 190 233 HEMOPEXIN-LIKE 3.
 FT DOMAIN 261 304 HEMOPEXIN-LIKE 4.
 FT DOMAIN 306 349 HEMOPEXIN-LIKE 5.
 FT METAL 81 81 IRON (HEME AXIAL LIGAND).
 FT METAL 152 152 IRON (HEME AXIAL LIGAND).
 FT METAL 152 152 IRON (HEME AXIAL LIGAND).
 FT DISULFID 52 233 BY SIMILARITY.
 FT DISULFID 151 156 BY SIMILARITY.
 FT DISULFID 190 202 BY SIMILARITY.
 FT DISULFID 255 458 BY SIMILARITY.
 FT DISULFID 364 406 BY SIMILARITY.
 FT DISULFID 416 433 BY SIMILARITY.
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 52 52 C -> W (IN REF. 2).
 FT HELIX 253 255
 FT TURN 257 258
 FT STRAND 263 266
 FT TURN 268 269
 FT STRAND 272 276
 FT TURN 277 278
 FT STRAND 282 282
 FT TURN 279 282
 FT STRAND 289 290
 FT STRAND 293 294
 FT STRAND 295 297
 FT HELIX 298 298
 FT TURN 300 301
 FT STRAND 308 312
 FT TURN 313 314
 FT STRAND 315 320
 FT TURN 321 322
 FT STRAND 323 328
 FT TURN 334 335
 FT STRAND 337 338
 FT TURN 341 342
 FT STRAND 343 347
 FT STRAND 360 362
 FT TURN 365 366
 FT STRAND 369 374
 FT TURN 375 376
 FT STRAND 377 382
 FT HELIX 383 388
 FT STRAND 392 394
 FT STRAND 403 407
 FT STRAND 422 427
 FT TURN 428 429
 FT STRAND 430 434
 FT HELIX 437 442
 FT STRAND 450 451
 FT HELIX 452 455
 FT TURN 456 457
 SQ SEQUENCE 460 AA; 5167 MW; 193BS985DEDF64EE CRC64;
 Query Match 77.8%; Score 1700.5; DB 1; Length 460;
 Best Local Similarity 67.7%; Pred. No. 2.6e-118;
 Matches 313; Conservative 26; Mismatches 50; Indels 73; Gaps 3;
 QY 1 MARVLGAPVALGLWSLCMSLAIAIPLPTPTSAHGNVAGE--TKDPDPTVTERCSDGWSFDA 58

Db 1 MVKASGIPIALGVGLWGLMSLATVNSVPLTSAHGNVTEGEGTKEADAVIIECCSDGWSFDA 60
 QY TTTLDNDGTMLEPFKGEFVWFKSHKMDRELISERWKNFPGPVDAAPFQGNNSVFLIKDKVMV 118
 Db 61 TTTLDNDGTMLEPFKGEFVWFKSHKMDRELISERWKNFPGPVDAAPFQGNNSVFLIKDKVMV 120
 QY 119 YPPEKKEKGYPKLQDFEPGIPSPPLDAAVECHRGCEQAEGLVFPQGH----- 165
 Db 121 YTSKNEKVPKSLQDFEPGIPSPPLDAAVECHRGCEQAEGLVFPQGNRKVMFDLTGTGKK 180
 QY 166 -----GHRN 169
 Db 181 ERSWPAVGNCTSALRWLGRRYCFQGNQFLRFNPVSGEVPBGYPLVDVDFLSCGRGHR- 239
 QY 170 GTGHGNSHTHGPEXWRCSPHVLVSALTSDNHGATYAFSGTHYVLDTSRDGMHSPYAHQ 229
 Db 240 -SSHRNSTQHGHESTRCDPDVLVSAMVSDNHGATYVSSGSHYRLDTRDGMHSPYAHQ 298
 QY 230 WPGPSAVDAAFSWEELVVOGTQVYVFLTKGGYTLVSGYPRLEKEVGTPHGIIDSV 289
 Db 299 WPGPSVYDAAFSWEELVVOGTQVYVFLTKGGYTLVSGYPRLEKEVGTPHGIIDSV 358
 QY 290 DAAPICGSSRLHIMAGRLVWLDLKSQAQATYTELFWPEKVDGALCMKSLGPNCSA 349
 Db 359 DAAPVCGSSRLHIMAGRLVWLDLKSQAQATYTELFWPEKVDGALCMKSLGPNCSA 418
 QY 350 NGPGLVLIHGNLYCYSDVEKLNAAKALPOPONTSLGGCTH 391
 Db 419 SGPNVLIHGNLYCYSDVEKLNAAKALPOPONTSLGGCTH 460
 RESULT 3
 HEMO RAT ID HEMO RAT STANDARD; PRT; 460 AA.
 AC P20059;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hemopexin precursor.
 GN HPX.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=91105180; Pubmed=1988069;
 RA Nikkila H., Gatliff J.D., Mueller-Eberhard U.;
 RT "Rat hemopexin. Molecular cloning, primary structural
 RT characterization, and analysis of gene expression.";
 RL Biochemistry 30:823-829 (1991).
 RN [2]
 RP SEQUENCE OF 24-53.
 RX MEDLINE=88339942; Pubmed=3421961;
 RA Wellner D., Cheng K.C., Mueller-Eberhard U.;
 RT "N-terminal amino acid sequences of the hemopexins from chicken, rat
 RT and rabbit.";
 RL Biochem. Biophys. Res. Commun. 155:622-625 (1988).
 RN [3]
 RP SEQUENCE OF 1-14 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=92287130; Pubmed=1599480;
 RA Nagae Y., Mueller-Eberhard U.;
 RT "Identification of an interleukin-6 responsive element and
 RT characterization of the proximal promoter region of the rat hemopexin
 RT gene.";
 RL Biochem. Biophys. Res. Commun. 185:420-429 (1992).
 CC -1- FUNCTION. BINDS HEME AND TRANSPORTS IT TO THE LIVER FOR BREAKDOWN
 CC AND IRON RECOVERY, AFTER WHICH THE FREE HEMOPEXIN RETURNS TO THE
 CC CIRCULATION.
 CC -1- SUBCELLULAR LOCATION: Secreted.

PN WO200166766-A2.
 XX 13-SEP-2001.
 PD 06-MAR-2001; 2001WO-US07167.
 XX PF 06-MAR-2000; 2000US-0187196.
 XX PR (DARM-) DARWIN MOLECULAR CORP.
 XX (SCHA/) SCHATZMAN R.
 PA
 PI Fajardo M, Wang K, Smith R, Moss P;
 XX WPI; 2001-582276/65.
 DR Novel isolated matrix metalloproteinase-25 nucleic acid molecule and
 XX proteins encoded by them whose inhibition is useful for modulation of
 PT hair growth in mammals -
 XX
 PS Example 2; Fig 3; 11pp; English.
 XX
 XX The present sequence is human matrix metalloproteinase (MMP)-12 protein
 CC used in the exemplification of the invention. MMP-25 DNA is located
 CC on chromosome 11q22. Matrix metalloproteinases are a family of zinc
 CC dependent endopeptidases that function extracellularly to degrade
 CC proteins typically found in the extracellular matrix. MMP-25 is expressed
 CC in skin cells of mammals, particularly in breast cells and hair
 CC follicles. MMP-25 DNA is useful for identifying a nucleic acid molecule
 CC encoding all or part of MMP by hybridizing MMP-25 to a nucleic acid
 CC sample and identifying a sequence that hybridizes in the nucleic
 CC acid sample. The identification step involves performing polymerase
 CC chain reaction (PCR) to amplify the hybridizing sequence. MMP-25
 CC antibody is useful for identifying type 25 MMP. MMP-25 protein
 CC inhibitors may be used to modulate hair growth and breast cancer in
 CC a mammal.
 CC
 XX Sequence 470 AA;
 SQ
 Query Match 9.3%; Score 202.5; DB 22; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;
 Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
 QY 30 SAHGVNAEGRTKDPDVT--RCSGMSPDATLDDNGTMLFFKGEFV---KSHKMDR 83
 DB 260 SLVGDPEKNQRLPNPNSEPALCDPNLSFDAVTVGN-KIFFKDRFFWLKVSERPTSV 318
 QY 84 ELISERKKNPSPVDAFR-QGHSVFLIKGDKVWVYPPPEKKGYPKLDQ-EFPGIPS 141
 DB 319 NLISSLMPTLPSCIEAAYIEARNQVFLFDKDKYMLSNLRPEPNPKSHSGFPPFVK 378
 QY 142 PLDAVECHRGECQAGVLFPOGHGRNGTGHGNSTHGPEYKCSPHLVLSALTSDNHG 201
 DB 379 KIDAAY-----FNRFYR----- 391
 QY 202 ATYAFSGTHYRLDTSRD-----GMHSWPIAHQMPQGSAYDAAFSWEK-LYLVQGTQVY 256
 DB 392 -TYFVDNQYWRDYDERQMDPGYPLKITKNFQIGPK-IDAVFYSKNKYYFFQGSNQP 449
 QY 257 VF 258
 DB 450 EY 451
 RESULT 7
 AAB84614
 ID AAB84614 standard; Protein; 470 AA.
 XX
 XX AAB84614;
 AC
 XX 05-SEP-2001 (first entry)
 DT
 XX Amino acid sequence of matrix metalloproteinase-12.
 DE
 XX

KM Growth factor; protein inhibitor; protease; damaged tissue;
 KM platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
 KM connective tissue derived growth factor; CTGF; chrysalin; YBGF;
 KM keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
 KM transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
 KM granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
 KM vascular endothelial growth factor; urokinase plasminogen activator;
 KM dermal ulcer; wound.
 XX
 XX Homo sapiens.
 OS
 PN WO200149309-A2.
 XX 12-JUL-2001.
 PD 21-DEC-2000; 2000MO-IB01935.
 XX 29-DEC-1999; 99GB-0030768.
 PR (PFI2) PFIZER LTD.
 PA (PFI2) PFIZER INC.
 XX
 PI Davies MJ, Huggins JP, McIntosh FS, Occleston NL;
 XX WPI; 2001-418351/44.
 DR N-PSDB; AAM28229.
 XX
 XX Composition for the treatment of damaged tissue i.e. chronic wounds and
 PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
 PT factor -
 XX
 PS Disclosure; Page 560; 572pp; English.
 XX
 XX The specification describes a pharmaceutical composition, comprising
 CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
 CC agent inhibits the action of at least one specific adverse protein,
 CC i.e. a protease, that is upregulated in a damaged tissue such as a
 CC wound environment. Growth factors which are included in the composition
 CC of the invention are platelet-derived growth factor (PDGF), fibroblast
 CC growth factor (FGF), connective tissue derived growth factor (CTGF),
 CC keratinocyte-derived growth factor (KGF), transforming factor (TGF),
 CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
 CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
 CC factor (VEGF), and chrysalin. Inhibitors which are included in the
 CC composition of the invention include inhibitors of urokinase-type
 CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
 CC composition is useful for the treatment of chronic damaged tissue, i.e.
 CC wounds and dermal ulcers. The present sequence represents a human MMP-12,
 CC and is used to produce the composition of the invention.
 CC
 XX Sequence 470 AA;
 SQ
 Query Match 9.3%; Score 202.5; DB 22; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;
 Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
 QY 30 SAHGVNAEGRTKDPDVT--RCSGMSPDATLDDNGTMLFFKGEFV---KSHKMDR 83
 DB 260 SLVGDPEKNQRLPNPNSEPALCDPNLSFDAVTVGN-KIFFKDRFFWLKVSERPTSV 318
 QY 84 ELISERKKNPSPVDAFR-QGHSVFLIKGDKVWVYPPPEKKGYPKLDQ-EFPGIPS 141
 DB 319 NLISSLMPTLPSCIEAAYIEARNQVFLFDKDKYMLSNLRPEPNPKSHSGFPPFVK 378
 QY 142 PLDAVECHRGECQAGVLFPOGHGRNGTGHGNSTHGPEYKCSPHLVLSALTSDNHG 201
 DB 379 KIDAAY-----FNRFYR----- 391
 QY 202 ATYAFSGTHYRLDTSRD-----GMHSWPIAHQMPQGSAYDAAFSWEK-LYLVQGTQVY 256
 DB 392 -TYFVDNQYWRDYDERQMDPGYPLKITKNFQIGPK-IDAVFYSKNKYYFFQGSNQP 449
 QY 257 VF 258

Db 450 EY 451

RESULT 8
AAB74595 standard; Protein; 470 AA.

AC AAB74595;

DT 11-JUN-2001 (first entry)

DE Human macrophage metalloelastase HME.

KW Human; HME: macrophage metalloelastase; elastolysis; pulmonary emphysema;
inflammatory disease.

OS Homo sapiens.

PN US6204043-B1.

PD 20-MAR-2001.

PF 01-MAR-1995; 95US-0396988.

PR 28-MAY-1993; 93US-0068392.

PA (UNIW) UNIV WASHINGTON.

PI Shapito SD;

DR WPI; 2001-289518/30.

DR N-PSDB; AAF81624.

PT New human macrophage metalloelastase, useful in the normal embryonic
development, growth, tissue remodeling and tissue repair, particularly
in studying the pathogenesis of pulmonary emphysema -

PS Claim 1; Fig 5; 20pp; English.

CC The present invention provides the protein and coding sequences of human
macrophage metalloelastase (HME). The protein has elastolytic activity.CC HME has a role in tissue remodeling and repair associated with
development and inflammation, and abnormal expression can result in
tumour invasiveness, arthritis and atherosclerosis. It is also thought to
contribute to the pathogenesis of pulmonary emphysema and other
inflammatory destructive diseases. The present sequence is the HME
protein.

CC Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 22; Length 470;

Best Local Similarity 27.3%; Pred. No. 3.9e-10; Mismatches 90; Indels 63; Gaps 10;

Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGETKDPDVT--RCSDEGSPDATTLDNGTMLFFKGEFVW----KSHKMDR 83

DB 260 SLVGDPKENGRLPNPNSEPALCDPNLSFDAVTTVGN-KIFFKDRFPLKVSERPTSV 318

QY 84 ELISERKKNFPPSVDAFR-QGHNSVFLIKDKRWVYPPKKEKGYPKLIQD-EFPGIPS 141

DB 319 NLISLWPTLPSSGIEAAYEIEARNQVFLFKDDKYWLISNLRPEPNPKSIHSFGFNFVK 378

QY 142 PLDAAYECHRGEQAGVLFQGHGHRNGTGHGNSTHHGPEVWRCSPHLVLSALTSNMG 201

DB 379 KIDAAV-----FNPRFYR----- 391

QY 202 ATYAFSGTHYRLDTSRD----GHSWPIAHQWPGPSAVDAAFSWEK-LYLVQGTQVY 256

DB 392 -TYFFVNDQYWRDERQMMDPGPKLITKNFGIGPK-IDAVFYKKNKYYFFQGSNQF 449

Db 450 EY 451

RESULT 9
AAB49982 standard; Protein; 470 AA.

AC AAB49982;

DT 13-MAR-2001 (first entry)

DE Human macrophage metalloelastase.

KW Human; macrophage metalloproteinase; elastin; matrix degrading enzyme;
emphysema.

OS Homo sapiens.

PN US6150152-A.

PD 21-NOV-2000.

PF 28-MAY-1993; 93US-0068392.

PR 28-MAY-1993; 93US-0068392.

PA (UNIW) UNIV WASHINGTON.

PI Shapito SD;

DR WPI; 2001-049090/06.

DR N-PSDB; AAC64994.

PT Novel human macrophage metalloelastase polynucleotides and polypeptides
useful for measuring elastin degradation -

PS Claim 1; Fig 5; 20pp; English.

CC The present invention provides the coding and protein sequences for the
human macrophage metalloelastase. This is a matrix degrading
metalloproteinase which has the ability to degrade elastin, and can be
used to measure elastin degradation.

CC Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 22; Length 470;

Best Local Similarity 27.3%; Pred. No. 3.9e-10; Mismatches 90; Indels 63; Gaps 10;

Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGETKDPDVT--RCSDEGSPDATTLDNGTMLFFKGEFVW----KSHKMDR 83

DB 260 SLVGDPKENGRLPNPNSEPALCDPNLSFDAVTTVGN-KIFFKDRFPLKVSERPTSV 318

QY 84 ELISERKKNFPPSVDAFR-QGHNSVFLIKDKRWVYPPKKEKGYPKLIQD-EFPGIPS 141

DB 319 NLISLWPTLPSSGIEAAYEIEARNQVFLFKDDKYWLISNLRPEPNPKSIHSFGFNFVK 378

QY 142 PLDAAYECHRGEQAGVLFQGHGHRNGTGHGNSTHHGPEVWRCSPHLVLSALTSNMG 201

DB 379 KIDAAV-----FNPRFYR----- 391

QY 202 ATYAFSGTHYRLDTSRD----GHSWPIAHQWPGPSAVDAAFSWEK-LYLVQGTQVY 256

DB 392 -TYFFVNDQYWRDERQMMDPGPKLITKNFGIGPK-IDAVFYKKNKYYFFQGSNQF 449

QY 257 VF 258

Db 450 EY 451

RESULT 10
AAB91061 standard; Protein; 470 AA.

XX AAU91061;
 AC 05-JUN-2002 (first entry)
 DT Human head and neck tumour protein, SEQ ID NO 219.
 DE Human; head and neck cancer; tumour; cytostatic; immunogenic; vaccine.
 KM Homo sapiens.
 XX WO200212329-A2.
 XX 14-FEB-2002.
 PD 01-AUG-2001; 2001WO-US24226.
 PF 03-AUG-2000; 2000US-223281P.
 XX 16-NOV-2000; 2000US-249933P.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Fan L;
 PI WPI; 2002-257467/30.
 DR N-PSDB; ABRK54023.
 XX Novel polynucleotide encoding head and neck tumour polypeptides, useful
 PT in pharmaceutical compositions, e.g. vaccines, for treating head and
 PS neck cancers -
 XX Claim 2; Page 181-182; 200P; English.
 CC The invention relates to an isolated polynucleotide (I) comprising
 CC sequences selected from 273 sequences fully defined in the specification.
 CC (I), including its encoded polypeptide (II), an antibody binding to (II),
 CC a fusion protein comprising (II) and a T-cell population stimulated by
 CC (I) or (II) are useful for stimulating an immune response in a patient
 CC and treating head and neck cancer in a patient. An oligonucleotide (III)
 CC that hybridises to (I) is useful for determining the presence of cancer
 CC in a patient, by obtaining a biological sample from the patient,
 CC contacting the sample with (III), detecting in the sample an amount of a
 CC polynucleotide that hybridises to the oligonucleotide, and comparing the
 CC amount of polynucleotide that hybridises to the oligonucleotide to a
 CC predetermined cut-off value. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis and
 CC treatment of head and neck cancer. AAU91061-AAU91062 represent human head
 CC and neck cancer protein sequences of the invention.
 CC
 SQ Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 23; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;

Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

30 SAHGNAVEGETPDDVTE--RCSGMSPDATLTDNGTMLPFKGEFVW-----KSHKMR 83
 DB 260 SLYGDKENQRLPNPNSEPALCDPNLSPDAVTTCN-KIFFKORFFWLKVSERKTSV 318
 QY 84 ELISERWKNFSPVDAAFR--QGHNSVFLIKGDKVWVYPPKKEKGYPKLLQD-EFGGIS 141
 DB 319 NLISSIMPTLPFGIEAAYEIBARNQVFLPKDDKVMYLSNLRBPNNPKSIHSGFPNFWK 378
 QY 142 PLDAVECHRGECQAEGLVFPQGHGRNGTGHGNSHTHGPEYRCSPHLVLSALTSDNHG 201
 DB 379 KIDAAV-----FNPFRYR----- 391
 QY 202 ATYAASGTHYWRDLTSRD---GMSWPIAHQWPOGSAVDAAFSWEK-LYLVQGTQVY 256
 DB 392 -TYFFVDNQYWRVDERKQMMDPGYPKLITKNFGIGPK-IDAVFYSKNKYYTFPGSSNPF 449
 QY 257 VP 258

DB 450 EX 451

RESULT 11

ABP96800
 ID ABP96800 standard; Protein; 470 AA.

AC ABP96800;

DT 05-JUN-2003 (first entry)

DE Human COPD related protein SEQ ID NO:50.

XX Human; chronic obstructive pulmonary disease; COPD; chronic lung disease.

XX Homo sapiens.

XX WO200297127-A2.

XX 05-DEC-2002.

XX 28-MAY-2002; 2002WO-EP05835.

XX 31-MAY-2001; 2001GB-0013266.

XX (FARB) BAYER AG.

XX Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;

XX WPI; 2003-140492/13.

XX N-PSDB; ACC46771.

XX Predicting, diagnosing or prognosing chronic lung disease, by detecting

XX a chronic obstructive pulmonary disease (COPD) gene in a biological

XX sample -

XX Claim 8; Page 153-155; 214P; English.

CC The present invention describes a method for predicting, diagnosing or
 CC prognosing chronic lung disease by detecting a chronic obstructive
 CC pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
 CC ACC46777, which encode the COPD related proteins in ABP96779 to
 CC ABP96806). The method is useful for predicting, diagnosing or prognosing
 CC chronic lung disease in a biological sample. The COPD genes and proteins
 CC encoded by them from the present invention (I) can be used for treating
 CC or preventing chronic lung disease in a mammal. (I) can be used in an
 CC animal model for determining the efficacy, toxicity, or side effects of
 CC treatment with (I), and determining the mechanism of action of (I).
 CC ACC46778 to ACC46903 represent COPD related PCR primers and probes used
 CC in an example from the present invention.
 CC
 SQ Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 24; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;

Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

30 SAHGNAVEGETPDDVTE--RCSGMSPDATLTDNGTMLPFKGEFVW-----KSHKMR 83
 DB 260 SLYGDKENQRLPNPNSEPALCDPNLSPDAVTTCN-KIFFKORFFWLKVSERKTSV 318
 QY 84 ELISERWKNFSPVDAAFR--QGHNSVFLIKGDKVWVYPPKKEKGYPKLLQD-EFGGIS 141
 DB 319 NLISSIMPTLPFGIEAAYEIBARNQVFLPKDDKVMYLSNLRBPNNPKSIHSGFPNFWK 378
 QY 142 PLDAVECHRGECQAEGLVFPQGHGRNGTGHGNSHTHGPEYRCSPHLVLSALTSDNHG 201
 DB 379 KIDAAV-----FNPFRYR----- 391
 QY 202 ATYAASGTHYWRDLTSRD---GMSWPIAHQWPOGSAVDAAFSWEK-LYLVQGTQVY 256
 DB 392 -TYFFVDNQYWRVDERKQMMDPGYPKLITKNFGIGPK-IDAVFYSKNKYYTFPGSSNPF 449

QY 257 VF 258
DB 450 EY 451

RESULT 12

ABR48203 standard; Protein; 470 AA.

ABR48203;

12-JUN-2003 (first entry)

Human bladder cancer associated protein sequence SEQ ID NO:125.

Human; bladder cancer; cytostatic; gene therapy; vaccine.

Homo sapiens.

WO2003003906-A2.

16-JAN-2003.

03-JUL-2002; 2002WO-US21338.

03-JUL-2001; 2001US-302814P.

03-AUG-2001; 2001US-310099P.

08-NOV-2001; 2001US-343705P.

13-NOV-2001; 2001US-350666P.

12-APR-2002; 2002US-372246P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Mack DH, Aziz N;

WPI: 2003-201532/19.

N-PSDB; ACC51017.

Claim 10, Page 275; 307pp; English.

The present invention describes a method for detecting a bladder cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence that is 80 % identical to a table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059 encode the human bladder cancer-associated sequences from the present invention (ABR48242). Bladder cancer-associated sequences from the present invention have cytostatic activities, and can be used in antisense gene therapy and in vaccine production. The method can be used for detecting a bladder cancer-associated transcript in a cell from a patient. The method is useful in diagnosing or treating bladder cancer and in screening for compounds that modulate bladder cancer, such as hormones or antibodies. The nucleic acid molecules from the present invention may be used in various screening and diagnostic methods, and for gene therapy, vaccine and/or antisense/inhibition applications.

Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 24; Length 470;

Best Local Similarity 27.3%; Pred. No. 3.9e-10; Mismatches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

30 SAHGNVAEGETKDDPVT--RSGDMSPTATLIDNGTMLFKGGEFV---KSHKMDR 83

260 SLVGDPKRNQRLNPNNSBPALCDPNISFPAVTVGN-KIFPFKDRFPMKVSEKPTSV 318

84 ELISERMKNRPSVDAAFR-QGHSNVLKSGDKRVVYPPPKKKGKGVKQLQD-EFPGIPS 141

319 NLISIMPTLPSGIEAAYEIANNOVFLFRDCKRWLISNLRPEPNYPKSIHSFGFPNFK 378

QY 142 PLDAVECHRGECQAGEVLFFQGHGRNGTGHGNSTHGPEYMRCSPHLVLSALTSBDMG 201

DB 379 KIDAAV-----FNRRFYR----- 391

QY 202 ATYAFSGTHTWRLDTSRD---GMHSPIAHQHPQPSAVDAAFSBEK-LYLVQGTQVY 256

DB 392 -TYFVDNQYWRDYDERROMMDPGYPLRLITGNFOGIGPK-IDAVFYKKNKYTFPGQSNQF 449

QY 257 VF 258

DB 450 EY 451

RESULT 13

ABU56421 standard; Protein; 470 AA.

ABU56421;

02-APR-2003 (first entry)

Lung cancer-associated polypeptide #14.

Lung cancer-associated polypeptide; cytostatic; emphysema; antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis; small cell lung cancer; benign lesion; precancerous lesion; bronchitis; chronic obstructive pulmonary disease; hyperensitively pneumonitis; interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.

WO200286443-A2.

31-OCT-2002.

18-APR-2002; 2002WO-US12476.

18-APR-2001; 2001US-284770P.

10-MAY-2001; 2001US-290492P.

09-NOV-2001; 2001US-339245P.

13-NOV-2001; 2001US-350666P.

29-NOV-2001; 2001US-354370P.

12-APR-2002; 2002US-372246P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Aziz N, Murray R;

WPI: 2003-093161/08.

N-PSDB; ABX76137.

Claim 27, Page 196; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hyperensitively pneumonitis, interstitial

CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
 CC polypeptides of the invention.

XX Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 24; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;
 Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNAVEGETPDDPVTE--RCSDEMSFDTATLDDNGTMLPFKGFVW----KSHKMR 83
 DB 260 SLYGDKENQRLPNPNSEPALCDPLSPDAVTYGN-KIFFKQRFPLKXSERKTSV 318
 QY 84 ELISERKNFPPSPVDAAFR-OGHNSVFLIKGDKWVYPEKKEKGYPKLLQD-EPPGIS 141
 DB 319 NLISSLMPTLPSSGIEAAYEIEARNOVFLPKDKYMLISLRPEPNPKSHSGFPNPK 378
 QY 142 PDDAVECHRGECQAEGLVFPQGHGRNGTGHGNSTHGPEYWRCSPHLVLSALTSDNHG 201
 DB 379 KIDAAV-----FNPRFYR----- 391
 QY 202 ATYAAGGTHYWRDLDSRD---GMHSWPIAHQWPGPSAVDAAFSWEK-LVLYOQTOY 256
 DB 392 -TYFPVDNQWYRDERQMDPQPLITKNFGIGPK-IDAVFYSKNKYTFPGGSNOF 449
 QY 257 VF 258
 DB 450 EX 451

RESULT 14
 ABUS6663
 ID ABUS6663 standard; Protein; 470 AA.

XX ABUS6663;

DT 02-APR-2003 (first entry)

XX Lung cancer-associated polypeptide #256.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;
 XX antiinflammatory; antiaesthetic; non-small cell lung cancer; atelectasis;
 XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

PF 18-APR-2002; 2002WO-US12476.

XX 18-APR-2001; 2001US-284770P.

XX 10-MAY-2001; 2001US-290492P.

XX 09-NOV-2001; 2001US-339245P.

XX 13-NOV-2001; 2001US-350666P.

XX 29-NOV-2001; 2001US-334370P.

XX 12-APR-2002; 2002US-372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI, 2003-093161/08.

XX N-PSDB; ABX76392.

XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer -

XX Claim 27; Page 385-386; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
 CC polypeptides of the invention.

XX Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 24; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;
 Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNAVEGETPDDPVTE--RCSDEMSFDTATLDDNGTMLPFKGFVW----KSHKMR 83
 DB 260 SLYGDKENQRLPNPNSEPALCDPLSPDAVTYGN-KIFFKQRFPLKXSERKTSV 318
 QY 84 ELISERKNFPPSPVDAAFR-OGHNSVFLIKGDKWVYPEKKEKGYPKLLQD-EPPGIS 141
 DB 319 NLISSLMPTLPSSGIEAAYEIEARNOVFLPKDKYMLISLRPEPNPKSHSGFPNPK 378
 QY 142 PDDAVECHRGECQAEGLVFPQGHGRNGTGHGNSTHGPEYWRCSPHLVLSALTSDNHG 201
 DB 379 KIDAAV-----FNPRFYR----- 391
 QY 202 ATYAAGGTHYWRDLDSRD---GMHSWPIAHQWPGPSAVDAAFSWEK-LVLYOQTOY 256
 DB 392 -TYFPVDNQWYRDERQMDPQPLITKNFGIGPK-IDAVFYSKNKYTFPGGSNOF 449
 QY 257 VF 258
 DB 450 EX 451

RESULT 15

XX AAB43772
 ID AAB43772 standard; Protein; 473 AA.

XX AAB43772;

DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO.1217.

XX Human, cancer associated gene; cancer antigen; detection; cancer;
 XX diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
 XX antidiabetic; antiaesthetic; antineumatic; antiarthritic; antiviral;
 XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 XX dermatologic; neuroprotective; thrombolytic; coagulant; nocotropic;
 XX vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;
 XX immune disorder; hematopoietic cell disorder; autoimmune disorder;
 XX allergic reaction; graft versus host disease; organ rejection;
 XX haemostatic; thrombolytic; cardiovascular disorder; infection;
 XX neurological disease; drug screening.

OS Homo sapiens.
XX
PN W0200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2000-587533/55.
DR N-PSDB; AAC77981.
XX
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
XX
PS Claim 11; Page 1848-1850; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnary; immunomodulator;
CC antidiabetic; antihypertensive; antineoplastic; antiallergic;
CC antiinflammatory; antithyroid; antidiabetic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC neoplastic; vasotropic; antiproliferative and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 473 AA;
XX
Query Match 9.3%; Score 202.5; DB 21; Length 473;
Best Local Similarity 27.3%; Pred. No. 4e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
QY 30 SAHGNVABGRTKPDVTE--RCSGMSPDATLTDNGTMLFFKGFVW---KSHKMDR 83
DB 263 SLYGDPKENQRLPNPNSSEPLCDPLSPDAVTVGN-KIFFKDRFPLKXSERPTSV 321
QY 84 ELISERKKNFSPVDNAFR-QGNSVFLIKGDKVWYPPPEKKEGYPKLIQD-EPPGIPS 141
DB 322 NLISSLMLPTLPSCIEAAYIEARNQVFLFPDDKYMLISNLRPNPKSIHSFGFPNFK 381
QY 142 PLDAAYECHRGEGQAGVLFPGHGRNNGHNSTHGPEYMRCSPHVLVSALTSDNHG 201
DB 382 KIDAAV-----FNRFYR----- 394
QY 202 ATYAFSGTHYWRDLTSDR---GMSWPIAHQWPOGSAVDAAFSWEK-LYLVQGTQVY 256
DB 395 -TYFFVDNQYWRDERQWMDPGYPLITKFGIGPK-IDAVFYSKNTYYFFQGSNPF 452
QY 257 VF 258
DB 453 EY 454

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:46:44 ; Search time 28 seconds
(without alignments)
590.841 Million cell updates/sec

Title: US-09-900-448-2

Perfect score: 2185

Sequence: 1 MARVLGAPVAGLWLSLWLSL.....NAKALPQPNQNTSLIGCTH 391

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6C_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	202.5	9.3	470	3	US-08-068-392-2
2	202.5	9.3	470	3	US-08-396-988-2
3	202.5	9.3	470	3	US-09-391-104-26
4	199	9.1	517	4	US-09-391-104-32
5	199	9.1	519	3	US-09-211-704-7
6	189.5	8.7	476	3	US-08-704-711A-21
7	189.5	8.7	476	3	US-08-448-489-14
8	189.5	8.7	476	3	US-09-521-220-21
9	189.5	8.7	476	4	US-09-391-104-22
10	183.5	8.4	444	1	US-09-178-002-2
11	183.5	8.4	467	4	US-09-391-104-24
12	183.5	8.4	467	4	US-08-448-489-13
13	183.5	8.4	468	3	US-08-448-489-13
14	181.5	8.3	469	4	US-09-391-104-23
15	181.5	8.3	492	1	US-07-794-393-4
16	181.5	8.3	492	1	US-08-001-711-4
17	179.5	8.2	488	1	US-07-794-393-2
18	179.5	8.2	488	1	US-08-001-711-2
19	179.5	8.2	488	3	US-08-704-711A-22
20	179.5	8.2	488	4	US-09-521-220-22
21	179.5	8.2	488	4	US-09-391-104-31
22	179.5	8.2	489	3	US-08-448-489-11
23	177	8.1	477	3	US-08-704-711A-20
24	177	8.1	477	3	US-08-448-489-15
25	177	8.1	477	3	US-08-281-313-1
26	177	8.1	477	4	US-09-521-220-20
27	177	8.1	477	4	US-09-391-104-21

28	176.5	8.1	469	3	US-08-704-711A-16	Sequence 16, Appl
29	176.5	8.1	469	3	US-08-448-489-12	Sequence 12, Appl
30	176.5	8.1	469	4	US-09-521-220-16	Sequence 16, Appl
31	176	8.1	471	4	US-09-391-104-25	Sequence 25, Appl
32	175	8.0	466	3	US-08-704-711A-17	Sequence 17, Appl
33	175	8.0	466	4	US-09-521-220-17	Sequence 17, Appl
34	172	7.9	564	3	US-09-211-704A-8	Sequence 8, Appl
35	172	7.9	669	3	US-08-704-711A-3	Sequence 3, Appl
36	172	7.9	669	4	US-09-521-220-3	Sequence 3, Appl
37	172	7.9	669	4	US-09-391-104-29	Sequence 29, Appl
38	169.5	7.8	508	4	US-09-391-104-18	Sequence 18, Appl
39	167.5	7.7	411	4	US-09-171-545-3	Sequence 3, Appl
40	167.5	7.7	416	4	US-09-171-545-4	Sequence 4, Appl
41	167.5	7.7	508	4	US-09-171-545-1	Sequence 1, Appl
42	164.5	7.5	324	2	US-08-816-755-2	Sequence 2, Appl
43	164.5	7.5	324	3	US-09-090-673-2	Sequence 2, Appl
44	161.5	7.4	532	3	US-09-294-841-2	Sequence 2, Appl
45	154	7.0	419	4	US-09-171-545-5	Sequence 5, Appl

ALIGNMENTS

```
RESULT 1
US-08-068-392-2
; Sequence 2, Application US/08068392
; Patent No. 6150152
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,392
; FILING DATE: 19930528
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25275
; REFERENCE/DOCKET NUMBER: 07-24(12406)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 694-3117
; TELEFAX: (314) 694-5435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-068-392-2

Query Match 9.3%; Score 202.5; DB 3; Length 470;
Best Local Similarity 27.3%; Pred. No. 2.4e-12;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGVAAGETPDVDYR--RCSDGMSPDATTLDDNGTMLPFKGBFV---KSHKMR 83
DB 260 SYVGPKENQRLPNPDNSPALCDPNLSFDATVTGN-KIFFKORFLKVSERPKSV 318
QY 84 ELISERKKNFSPVAAFR-QGHSNVFLIKGDVWVYPEKKEKGYPKLQD-EFGGIPS 141
DB 319 NLISSLMPTLPGLIAAYIABRNQVFLFKDKKWLISLRLREPRYPKSHSFGFPNFK 378
```

QY	142	PLDAVGECHRGCEQAEVGLFPGGHGRNGTGHGNTGHGPEYMRCSPHVLVALTSNDNG	201
DB	379	KIDAAV-----FNPRYR-----	391
QY	202	ATVAFSGTHWRRLDTSRD---GHWSPILAHQWPGQPSAVDAFAFSWEK-LYLVQGTQYV	256
DB	392	-TYFVFDNQYRYDERBQRQMDPGYFKLITKNFGQIGPK-IDAVFYSKNXYFFQGSNGF	449
QY	257	VF 258	
DB	450	EY 451	
RESULT 2			
US	08-396-988-2	Sequence 2, Application US/08396988	
		Patent No. 6204043	
		GENERAL INFORMATION:	
		APPLICANT: Shapiro, Steven M.	
		TITLE OF INVENTION: Human Macrophage Metalloproteinase	
		NUMBER OF SEQUENCES: 3	
		CORRESPONDENCE ADDRESS:	
		ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM	
		STREET: 800 N. Lindbergh Blvd.	
		CITY: St. Louis	
		STATE: Missouri	
		COUNTRY: USA	
		ZIP: 63167	
		COMPUTER READABLE FORM:	
		MEDIUM TYPE: Floppy disk	
		COMPUTER: IBM PC compatible	
		OPERATING SYSTEM: PC-DOS/MS-DOS	
		SOFTWARE: Patent Release #1.0, Version #1.25	
		CURRENT APPLICATION DATA:	
		APPLICATION NUMBER: US/08/396, 988	
		FILING DATE: 01-MAR-1995	
		CLASSIFICATION: 435	
		PRIOR APPLICATION DATA:	
		APPLICATION NUMBER: US 08/068,392	
		FILING DATE: 28-MAY-1993	
		ATTORNEY/AGENT INFORMATION:	
		NAME: Meyer, Scott J.	
		REGISTRATION NUMBER: 25275	
		REFERENCE/DOCKET NUMBER: 07-24(12406)A	
		TELECOMMUNICATION INFORMATION:	
		TELEPHONE: (314)694-3117	
		TELEFAX: (314)694-5435	
		INFORMATION FOR SEQ ID NO: 2:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 470 amino acids	
		TYPE: amino acid	
		TOPOLOGY: linear	
		MOLECULE TYPE: protein	
		US-08-396-988-2	
		Query Match	9.3%; Score 202.5; DB 3; Length 470;
		Best Local Similarity	27.3%; Pred. No. 2.4e-12;
		Matches	66; Conservative 23; Mismatches 90; Indels 63; Gaps 10
QY	30	SAHGNVAGETKRPDVTYE--RCSGWSFDDATLTDNGTMTLFFKGEFVW---KSHKMR	83
DB	260	SLYGDGPKENQRLPNPDNSEPALCDNLSPDVAITVGN-KIFFEKRFPWLKYSERPKTSV	318
QY	84	ELISERWKQNPSPVDAAFR-QGHSNVFLIKGDKWVYPPKKKEKYPKLQD-EFFGITS	141
DB	319	NLISSSLWFTLLSGIEAAVEIERKNVFLFKDKKYLISNLAREPNPKSIHSFGFNNFYK	378
QY	142	PLDAVGECHRGCEQAEVGLFPGGHGRNGTGHGNTGHGPEYMRCSPHVLVALTSNDNG	201
DB	379	KIDAAV-----FNPRYR-----	391
QY	202	ATVAFSGTHWRRLDTSRD---GHWSPILAHQWPGQPSAVDAFAFSWEK-LYLVQGTQYV	256

```

Db      392 -TYFVNQWRYRDERQMDDPGYPLKLTKNFGSIGPR-IDAIFYSKKKYYFFQGSGNQF 449
QY      257 VF 258
Db      450 EY 451

RESULT 3
US-09-391-104-26
; Sequence 26, Application US/09391104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; TITLE OF INVENTION: OF USING SAME
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 470
; TYPE: PRt
; ORGANISM: Homo sapiens
US-09-391-104-26

Query Match          9.3%; Score 202.5; DB 4; Length 470;
Best Local Similarity 27.3%; Pred.No. 2.4e-12;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10.

QY      30 SAGGNVABGETKPPDYTE--RCSDGMSFDAATTLDNGTMLPFKGEPFW----KSHKDR 83
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      260 SLVDPDPMENQRLPMPDBDSBPALCDPNLSFDVAITVGN-KIFFFDKRFMLKVSRPKTSV 318
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      84 ELISERKKNPSPVDAAFR-QGHNSVFLLIKDKYVVYPPEKKEGGYPLLD-EPFGIIPS 141
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      319 NLISLMLPTLDPSGLEAYVEIFARNOVELFKODKYMLSLNLPEEPPYPSISFGFPNVK 378
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      142 PLDAAVECHRGCEAEGVLFPQGHGRNGTGHSNTHHGPEVMCSPHLVLSALTSDNHG 201
       |||||-----FNPRFYR----- 391
Db      379 KIDAAV-----FNPRFYR----- 391

QY      202 ATYAFCSTHTWRDLTSD---GHSNPFIHQWFOGSANDAAASWEK-LYLVOGTQVY 256
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      392 -TYFVNQWRYRDERQMDDPGYPLKLTKNFGSIGPR-IDAIFYSKKKYYFFQGSGNQF 449
QY      257 VF 258
Db      450 EY 451

RESULT 4
US-09-391-104-32
; Sequence 32, Application US/09391104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07

```

PRIOR APPLICATION NUMBER: US 08/814,394
 PRIOR FILING DATE: 1997-03-11
 NUMBER OF SEQ ID NOS: 35
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 32
 LENGTH: 517
 TYPE: PRF
 ORGANISM: Homo sapiens
 US-09-391-104-32

Query Match 9.1%; Score 199; DB 4; Length 517;
 Best Local Similarity 25.0%; Pred. No. 6.3e-12;
 Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

QY 5 LGAPVALGL-----NSLCMSLAATPL-----PPTSAHGNAVAGETKPPDPVTERCS 51
 DB 188 VDDPLRYGRLPYEDKVRWOLYGVRESVSPTAQPBEPPPLPEPPDNRSSAPPRKDVPHRCS 247
 QY 52 DGMSPDATTLDNGTMLPFKGEFVWKS HKMDRELISER-----WKNP-----SPVDAAF 101
 DB 248 T--HFDVA-VQIRGEAFPFKGYFWRLTR-DRLVLSLQPAQMRFRGLPLHDSVDAY 303
 QY 102 -RQHSNVELIGDKVWYPPPEKKEGPKLQDEFPPIPSPLDAAVECHRGCEQAEVYL 160
 DB 304 ERTSDHKIVFFKGDRTWVFKDNNVEGYRPSVDFSLPPGIDAA----- 348
 QY 161 FQGHGRNGTGHGNSHTHGPEYMRCSPHLVLSALTSNDHGATYAFSGTHYRLDT-SRD 219
 DB 349 FSWAHNDR-----TYFFKQDLVYMYDDHTKH 374
 QY 220 GHSWPI-AHQWPGSVAVDAAFSWEE-KLYLVQGTQVYVFLTKGGYTLVSGYKLEKE 277
 DB 375 MDGPYAGSPLRWGVSTLDDAMRMSDGASYFRG-QEYWKVLDELVAAPGPOSTARD 433
 QY 278 -----VGTPHGIILDSVDAA 292
 DB 434 WLVCGDSDQADGSVAAGVDAA 453

RESULT 5
 US-09-211-704A-7
 Sequence 7, Application US/09211704A
 Patent No. 6271014
 GENERAL INFORMATION:
 APPLICANT: de Saint-Vie, Blandine Marie
 APPLICANT: Fossiez, Francois
 APPLICANT: Caux, Christophe
 APPLICANT: Lebeque, Serge J.B.
 TITLE OF INVENTION: Mammalian Proteases; Related Reagents
 TITLE OF INVENTION: and Methods
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DNAX Research Institute
 STREET: 901 California Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/211,704A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/005,263
 FILING DATE: 09-JAN-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Ching, Edwin P.
 REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: SF0781K
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650)852-9196
 TELEFAX: (650)496-1200
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 519 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-211-704A-7

Query Match 9.1%; Score 199; DB 3; Length 519;
 Best Local Similarity 25.0%; Pred. No. 6.4e-12;
 Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

QY 5 LGAPVALGL-----NSLCMSLAATPL-----PPTSAHGNAVAGETKPPDPVTERCS 51
 DB 190 VDDPLRYGRLPYEDKVRWOLYGVRESVSPTAQPBEPPPLPEPPDNRSSAPPRKDVPHRCS 249
 QY 52 DGMSPDATTLDNGTMLPFKGEFVWKS HKMDRELISER-----WKNP-----SPVDAAF 101
 DB 250 T--HFDVA-VQIRGEAFPFKGYFWRLTR-DRLVLSLQPAQMRFRGLPLHDSVDAY 305
 QY 102 -RQHSNVELIGDKVWYPPPEKKEGPKLQDEFPPIPSPLDAAVECHRGCEQAEVYL 160
 DB 306 ERTSDHKIVFFKGDRTWVFKDNNVEGYRPSVDFSLPPGIDAA----- 350
 QY 161 FQGHGRNGTGHGNSHTHGPEYMRCSPHLVLSALTSNDHGATYAFSGTHYRLDT-SRD 219
 DB 351 FSWAHNDR-----TYFFKQDLVYMYDDHTKH 376
 QY 220 GHSWPI-AHQWPGSVAVDAAFSWEE-KLYLVQGTQVYVFLTKGGYTLVSGYKLEKE 277
 DB 377 MDGPYAGSPLRWGVSTLDDAMRMSDGASYFRG-QEYWKVLDELVAAPGPOSTARD 435
 QY 278 -----VGTPHGIILDSVDAA 292
 DB 436 WLVCGDSDQADGSVAAGVDAA 455

RESULT 6
 US-08-704-711A-21
 Sequence 21, Application US/08704711A
 Patent No. 6114159
 GENERAL INFORMATION:
 APPLICANT: WILLY, Horst
 APPLICANT: HINZMANN, Bernd
 TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
 TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/704,711A
 FILING DATE: 20-NOV-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/DE95/00357
 FILING DATE: 17-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 4438838.1

FILING DATE: 21-OCT-1994
PRIOR APPLICATION DATA: DE 4409663.1
APPLICATION NUMBER: 17-MAR-1994
FILING DATE: 17-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 26083/124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-704-711A-21

Query Match 8.7%; Score 189.5; DB 3; Length 476;
Best Local Similarity 28.4%; Pred. No. 5.4e-11;

Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

QY 25 PLPTSAHGVABGETKPDVTERCSDDGWSFDA-TTLDNGTMTLFFKGEFVWKSHKMDR 83
DB 272 PLVPTK---SVPSGSEMP-----AKCDPALSPDAISTL--RGEYLFKORYFWRSHMNP 321

QY 84 E-----LISERWKNFSPSPVDAAFR-QGHNVSFLIKGDKVWV-----YP 120

DB 322 EPEFHLSAFWPSLPYLDAAYEVNSRDYVFIKGNFMAIRGNEVOAGYPRGHTLGF 381

QY 121 P-----EKK-----EKGPKLQDEFPGIPSPDLAAVE 148
DB 382 PTIRKIDAAVSDKEKKTYFFADKYWRPDENSQSGEPRLIADDPGEVPRKVAVL- 440

QY 149 CHRGECAQGVLPF 162
DB 441 -----QAFGFYF 448

RESULT 7
US-08-448-489-14
Sequence 14, Application US/08448489
Patent No. 6184022
GENERAL INFORMATION:
APPLICANT: SEIKI, Mochiharu
APPLICANT: SATO, Hiroshi
APPLICANT: SHINAGAWA, Akira
TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
FILE REFERENCE: 55-290P
CURRENT FILING DATE: US/08/448, 489
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 476
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Known Member of
OTHER INFORMATION: Matrix Metalloproteinase Family
US-08-448-489-14

Query Match 8.7%; Score 189.5; DB 3; Length 476;
Best Local Similarity 28.4%; Pred. No. 5.4e-11;
Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

QY 25 PLPTSAHGVABGETKPDVTERCSDDGWSFDA-TTLDNGTMTLFFKGEFVWKSHKMDR 83
DB 272 PLVPTK---SVPSGSEMP-----AKCDPALSPDAISTL--RGEYLFKORYFWRSHMNP 321
QY 84 E-----LISERWKNFSPSPVDAAFR-QGHNVSFLIKGDKVWV-----YP 120

DB 322 EPEFHLSAFWPSLPYLDAAYEVNSRDYVFIKGNFMAIRGNEVOAGYPRGHTLGF 381
QY 121 P-----EKK-----EKGPKLQDEFPGIPSPDLAAVE 148
DB 382 PTIRKIDAAVSDKEKKTYFFADKYWRPDENSQSGEPRLIADDPGEVPRKVAVL- 440
QY 149 CHRGECAQGVLPF 162
DB 441 -----QAFGFYF 448

RESULT 8
US-09-521-220-21
Sequence 21, Application US/09521220
Patent No. 6393348
GENERAL INFORMATION:
APPLICANT: WILF, Horet
HINZMANN, Bernd
TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
METALLOPROTEASES, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/521,220
FILING DATE: 08-Mar-2000
CLASSIFICATION: <Unknown>
21-OCT-1994
17-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/704,711
FILING DATE: <Unknown>
APPLICATION NUMBER: DE 4438838.1
FILING DATE: 21-OCT-1994
APPLICATION NUMBER: DE 4409663.1
FILING DATE: 17-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 26083/124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-521-220-21

Query Match 8.7%; Score 189.5; DB 4; Length 476;
Best Local Similarity 28.4%; Pred. No. 5.4e-11;
Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

QY 25 PLPTSAHGVABGETKPDVTERCSDDGWSFDA-TTLDNGTMTLFFKGEFVWKSHKMDR 83
DB 272 PLVPTK---SVPSGSEMP-----AKCDPALSPDAISTL--RGEYLFKORYFWRSHMNP 321
QY 84 E-----LISERWKNFSPSPVDAAFR-QGHNVSFLIKGDKVWV-----YP 120

```

; ORGANISM: Homo sapiens
; US-09-178-002-2
;
; Query Match
; Best Local Similarity 28.4%; Pred. No. 5,4e-11;
; Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;
;
; QY 25 PLPTSAHGNVABGETKPDVTERCSGDSGSPDA-TTLDNGTMLPFKGEFVWKSMDR 83
; DB 322 EPEFHLSAFWPSLPSTLDAAYEVNSRDYVIFKGNFMAIRGNEVOAGYPRGIHTLGF 381
; QY 121 P-----EKK-----EKGYPKLQDEFFGIPSPDLAAVE 148
; DB 382 PTIRKIDAAVSDKEKKTYFFAADKYWRPDENSQSGFPRLIADFPGEVEKVDVAVL- 440
; QY 149 CHRGECCOAGSVLFF 162
; DB 441 -----QARGFFYF 448
;
; RESULT 9
; US-09-391-104-22
; Sequence 22, Application US/09391104
; Patent No. 639371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; PRIOR FILING DATE: 1999-09-07
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-391-104-22
;
; Query Match
; Best Local Similarity 28.4%; Pred. No. 5,4e-11;
; Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;
;
; QY 25 PLPTSAHGNVABGETKPDVTERCSGDSGSPDA-TTLDNGTMLPFKGEFVWKSMDR 83
; DB 322 EPEFHLSAFWPSLPSTLDAAYEVNSRDYVIFKGNFMAIRGNEVOAGYPRGIHTLGF 381
; QY 121 P-----EKK-----EKGYPKLQDEFFGIPSPDLAAVE 148
; DB 382 PTIRKIDAAVSDKEKKTYFFAADKYWRPDENSQSGFPRLIADFPGEVEKVDVAVL- 440
; QY 149 CHRGECCOAGSVLFF 162
; DB 441 -----QARGFFYF 448
;
; RESULT 10
; US-09-178-002-2
; Sequence 2, Application US/09178002
; Patent No. H001973
; GENERAL INFORMATION:
; APPLICANT: Hu, Shou-Ih
; TITLE OF INVENTION: Human Neutrophil Collagenase Splice Variant
; FILE REFERENCE: CGC 2048
; CURRENT APPLICATION NUMBER: US/09/178,002
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 444
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
; US-09-178-002-2
;
; Query Match
; Best Local Similarity 24.4%; Pred. No. 2e-10;
; Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;
;
; QY 23 ATPPTSAHGNVABGETKPDVTERCSGDSGSPDA-TTLDNGTMLPFKGEFVWKSMDR 79
; DB 242 SNDIQPT-----GPSTPKP-----CDPSLTFDAITTL--RGEILFFKDRYFMRHPQ 286
; QY 80 --KMDRELISERKKNPSPVDAAFRQ-GHNSVFLIKGDKVWVYPPEKGEYKLLQD-B 135
; DB 287 LQREVNPFISLFWPSLPSTLDAAYEDPDRDLFLFKGNQYMLSGVDILQGYPKDISNYG 346
; QY 136 FPGIPSPDLAAVECHRGECOAGSVLFFQGHGHRNGTGHGNSTHGPEYWRCSPLVLVAL 195
; DB 347 FPESSVAIDAAV-----FYR----- 361
; QY 196 TSDNHGATYAFSGTHYWRDLTSR---DGMHSWPIAHQWPGSAVDAAFSWEKLYLVQG 252
; DB 362 ----SKTYFVNDDQFWRVNDQRFWMEPGYPSISGAPEISKVDVAVFOQEHFFVFSG 416
; QY 253 TQYVYF 258
; DB 417 PRYIAF 422
;
; RESULT 11
; US-09-178-002-4
; Sequence 4, Application US/09178002
; Patent No. H001973
; GENERAL INFORMATION:
; APPLICANT: Hu, Shou-Ih
; TITLE OF INVENTION: Human Neutrophil Collagenase Splice Variant
; FILE REFERENCE: CGC 2048
; CURRENT APPLICATION NUMBER: US/09/178,002
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-178-002-4
;
; Query Match
; Best Local Similarity 24.4%; Pred. No. 2,2e-10;
; Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;
;
; QY 23 ATPPTSAHGNVABGETKPDVTERCSGDSGSPDA-TTLDNGTMLPFKGEFVWKSMDR 79
; DB 265 SNDIQPT-----GPSTPKP-----CDPSLTFDAITTL--RGEILFFKDRYFMRHPQ 309
; QY 80 --KMDRELISERKKNPSPVDAAFRQ-GHNSVFLIKGDKVWVYPPEKGEYKLLQD-B 135
; DB 310 LQREVNPFISLFWPSLPSTLDAAYEDPDRDLFLFKGNQYMLSGVDILQGYPKDISNYG 369
; QY 136 FPGIPSPDLAAVECHRGECOAGSVLFFQGHGHRNGTGHGNSTHGPEYWRCSPLVLVAL 195
; DB 370 FPESSVAIDAAV-----FYR----- 384
; QY 196 TSDNHGATYAFSGTHYWRDLTSR---DGMHSWPIAHQWPGSAVDAAFSWEKLYLVQG 252
; DB 385 ----SKTYFVNDDQFWRVNDQRFWMEPGYPSISGAPEISKVDVAVFOQEHFFVFSG 439
; QY 253 TQYVYF 258
; DB 440 PRYIAF 445
;
; RESULT 12
; US-09-391-104-24

```

```

; Sequence 24, Application US/09391104
; Patent No. 6393371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Palduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-391-104-24

```

```

Query Match      8.4%; Score 183.5; DB 4; Length 467;
Best Local Similarity 24.4%; Pred. No. 2.2e-10;
Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;

```

```

QY 23 ATPLPPTSAHGNVAEGTKEPDVTERCSGWSFDA-TTLDNGTMLFFKGGEVWVSH-- 79
DB 265 SNPIQPT-----GSPSTPKP-----CDPSLTFDAITTL--RGEILFFKDRFWMRRHPQ 309
QY 80 --KMDSELISERKNPSPVDAAFRO-GHNSVFLIKGDKWVVPPEKKEGYKLLQD-E 135
DB 310 LQREVNPFISLFWPSLPTGIQAAYEDPRDLIFLFGNQWALSGYDILQGYKDISNYG 369
QY 136 FPGISPDLAAVECHRGECQAEGLVFFQGHGHRNGTGHGNSHGPYRMCSPHLVLSAL 195
DB 370 FPGSSVQALDAV-----FYR----- 384
QY 196 TSDNHGATYAFSGTHWRDLTSR---DGMHSPRIAHQWPGSPAVDAAFSWEKLYVQ 252
DB 385 -----SKTYFFVNDQFWRDYNQRFMEPGYPKISGAFPGIESKVDVAFQOEHPHFVFSG 439
QY 253 TVYVVF 258
DB 440 PRYIAF 445

```

```

RESULT 13
US-08-448-489-13

```

```

; Sequence 13, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: X = UNKNOWN
; OTHER INFORMATION: Description of Unknown Organism: Known Member of
; OTHER INFORMATION: Matrix Metalloproteinase Family
; US-08-448-489-13

```

```

Query Match      8.4%; Score 183.5; DB 3; Length 468;
Best Local Similarity 24.4%; Pred. No. 2.2e-10;
Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;

```

```

QY 23 ATPLPPTSAHGNVAEGTKEPDVTERCSGWSFDA-TTLDNGTMLFFKGGEVWVSH-- 79
DB 265 SNPIQPT-----GSPSTPKP-----CDPSLTFDAITTL--RGEILFFKDRFWMRRHPQ 309
QY 80 --KMDSELISERKNPSPVDAAFRO-GHNSVFLIKGDKWVVPPEKKEGYKLLQD-E 135
DB 310 LQREVNPFISLFWPSLPTGIQAAYEDPRDLIFLFGNQWALSGYDILQGYKDISNYG 369
QY 136 FPGISPDLAAVECHRGECQAEGLVFFQGHGHRNGTGHGNSHGPYRMCSPHLVLSAL 195
DB 370 FPGSSVQALDAV-----FYR----- 384
QY 196 TSDNHGATYAFSGTHWRDLTSR---DGMHSPRIAHQWPGSPAVDAAFSWEKLYVQ 252
DB 385 -----SKTYFFVNDQFWRDYNQRFMEPGYPKISGAFPGIESKVDVAFQOEHPHFVFSG 439
QY 253 TVYVVF 258
DB 440 PRYIAF 445

```

```

RESULT 14
US-09-391-104-23

```

```

; Sequence 23, Application US/09391104
; Patent No. 6393371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Palduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-391-104-23

```

```

Query Match      8.3%; Score 181.5; DB 4; Length 469;
Best Local Similarity 26.3%; Pred. No. 3.6e-10;
Matches 59; Conservative 26; Mismatches 82; Indels 57; Gaps 8;

```

```

QY 44 PVTERCSGWSFDA-TTLDNGTMLFFKGGEVWVSHKDRDL-----ISERKNPSPVD 98
DB 272 PGPACDCKLTFDAITTL--RGEVWFKDRFYMRNTPPYKELNFIIVFNPQLNGE 329
QY 99 AAFR-QHNSVFLIKGDKWVVPPEKKEGYKLLQDEPFGIPSPDLAAVECHRGECQAE 157
DB 330 AAYEPADRDEVPFFPKNKYMAVOGQVNLHGYPKDYISF-GPP----- 371
QY 158 GVLFFQGHGHRNGTGHGNSHGPYRMCSPHLVLSALSDNHGATYAFSGTHWRDL-- 215
DB 372 -----RTVKA-IDAAUSEENTGTYFFVANKTMYDEY 403
QY 216 -TSRDGHSMPRIAHQWPGSPAVDAAFSWEKLYVQTVYVVF 258
DB 404 KXSMDEGPKMTAHDPFGIGHVDAVFMKDGFFYPFHGRQYKF 447

```

```

RESULT 15
US-07-794-393-4

```



```

; Sequence 4, Application US/07794393
; Patent No. 5236844
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, PIERRE
; APPLICANT: BASSET, PAUL
; APPLICANT: BELLOU, JEAN-PIERRE
; TITLE OF INVENTION: ANALYTICAL MARKERS FOR MALIGNANT BREAST
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/794,393
; FILING DATE: 19911121
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9025326.1
; FILING DATE: 21-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1383.0040000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 492 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-794-393-4

Query Match      8.3%; Score 181.5; DB 1; Length 492;
Best Local Similarity 25.2%; Pred. No. 3.8e-10;
Matches 69; Conservative 26; Mismatches 86; Indels 93; Gaps 13;

QY 24 TELPPTSAH-----GNVAGETKPDVTERCSGWSFDATLDDNGTMLFFKGEFVWKS 78
DB 270 SPAPTISSQAGDTNINIALEPETPPDVCE-----SFDVNS-TIRGELFFKAGFVWRL 323
QY 79 HKMDRE-----LISRWKNPSPFVDAAFROGHSVFLIKGDKVWVYPPEKEKGYPKLIQ 133
DB 324 RSGRLQPGYPALASRMWGLPSFVDAAFDAQOIWFQQAQYWVYDGEKPVLGPAFLSK 383
QY 134 DEFPGLPSPLDAVECHRECCQAEGLVLFQGHGRNGTGHGNSTHGPEYMRCSPHILVS 193
DB 384 IGLIQG--SPVHALV-----WGPR-----400
QY 194 ALTSDNHGATYAFSGTHYWRLD--TSR-DGWSHWPILAHQ--WPGPSAVDAAFSWEKTL 247
DB 401 -----KNKIYFFRGADYWRFHPRTRQVDN-----PVRKSTDMRGVPSSEIDAAF-----444
QY 248 YLVQGTQYVYVFLTKG-----GYTLVSGYPK 272
DB 445 ---QDAEGYAYFLRGHLWYKFDVVKVYLGFPFR 475

```

Search completed: December 16, 2003, 06:55:24
 Job time : 29 secs

THIS PAGE BLANK (USPTO)

```

121 PEEKKEKGPXKLDPEFPGPSPDLDAVECHRGECOEGULFFQ----- 163
|||||

```

Db 121 PEKKEKGYPKLQDEPFPGIPSPDLAAVECHRGECQAGVLPFGQDREMFMDLATGTWKER 180
 QY 164 -----GCHGRN 169
 Db 181 SWPVGNCSSALRWLGRYCGQGNQFLRDPVGEVPPRRYPRDYVDFMPCPGHGRN 240
 QY 170 GTGHGNS 176
 241 GTGHGNS 247

RESULT 2

Q8KUI6 PRELIMINARY; PRT; 276 AA.
 AC Q8KUI6; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DE Hemopexin (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Alam J., Smith A.;
 RT "Nucleotide sequence of the mouse haemopexin gene";
 RL Submitted (DEC-1990) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X56829; CAA40160.1; -
 DR Interpro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 3.
 DR SMART; SM00120; HX; 3.
 FT NON TER 276
 SQ SEQUENCE 276 AA; 31047 MW; BD336C649B2B767 CRC64;

Query Match 31.6%; Score 689.5; DB 11; Length 276;
 Best Local Similarity 53.0%; Pred. No. 7.8e-50;
 Matches 149; Conservative 12; Mismatches 41; Indels 79; Gaps 6;

QY 1 MARVLGAPVALGLMSLCWSLAATPRLPTSAHGNVAEGE--TKRDPVTERCSQSGWSPDA 58
 Db 1 MARTAVLNLIVLGLCWSLAVASPLP--IAHGVALEVNGTKDSDVPERCLDTWGFDA 58
 QY 59 TLLDDNGTMLPFKGEFVWKSHKMDRELISRWKKNFSPVDAAPFGQNSVFLIKGDVWV 118
 Db 59 ATMDHNGTMLPFKGEFVWKSHKMDRELISRWKKNFSPVDAAPFGQNSVFLIKEDVWV 117
 QY 119 YPEKKEKGYPKLQDEPFPGIPSPDLAAVECHRGECQAGVLPFGQ----- 165
 Db 118 YPEKKEKGYPKLQDEPFPGIPSPDLAAVECHRGECQAGVLPFGQNRKKVFMDFATRTQK 177
 QY 166 -----GH-- 167
 Db 178 ERSGPAVGNCTAALRWLEERYCFQGNKFLRFNPVTGEVPPRRYPRDYVDFMPCPGHGR 237
 QY 168 -RNGTGHGNSHTHGRPEWRKCSPHLVLSALSDNKGATYAPS 207
 Db 238 PRNGTGHGNSHTHGRPEWRKCSPHLVLSALSDNKGATYAPT 276

RESULT 3

Q8UIP8 PRELIMINARY; PRT; 427 AA.
 AC Q8UIP8; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DE Warm-temperature-acclimation-related-65 kDa-protein-like-protein.
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;

OC Belontiiformes; Adriantichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=HNI; TISSUE=Liver;
 RA Hirayama M., Kobiyama A., Kinoshita S., Watabe S.;
 RT "Primary structural analysis of medaka Wap65 and Wap65-like protein
 RT and their gene expression in association with temperature acclimation
 RT and during ontogeny";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB075199; BAB97304.1; -
 DR Interpro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 3.
 DR SMART; SM00120; HX; 7.
 SQ SEQUENCE 427 AA; 48000 MW; DED6B6A5BND38EB CRC64;

Query Match 28.6%; Score 625; DB 13; Length 427;
 Best Local Similarity 33.0%; Pred. No. 3.5e-44;
 Matches 146; Conservative 66; Mismatches 136; Indels 94; Gaps 12;

QY 15 SLCWSLAATPRLPTSAHGNVAEGEIK-PDPVTERCSQSGWSPDATLDDNGTMLPFKGE 73
 Db 11 ALMLALTRAAPLEDSAA---GDGSAALPRGCA-----GIEFDATTPDDKQGFPEKGD 60
 QY 74 FVWKSHKMDRELISRWKKNFSPVDAAPR-----QG-HNSVFLIKGDVWVYPPEKKEK 126
 Db 61 HWKKGEGDAQPSQYFKELNGHVDAAPRMHNPENQGDHDIYLFDDKVFSPFHTLBE 120
 QY 127 GYPKLLQDEPFPGIPSPDLAAVECHRGECQAGVLPFGQ----- 164
 Db 121 GYPKELQDEPFPGIPSPDLAAVECHRGECQAGVLPFGQDVHVVYDITVTKVTKTPHLP 180
 QY 165 -----HGR-----NGTGHGNSHTHGRPE--MRC-----SP 188
 Db 181 ACTSVRMLEHYCYCFGHNFRFPQVYTGVTGYPVDAARFYERKCPDFHGERRTLKCS 240
 QY 189 HLVLALTSQDNHGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 248
 Db 241 DFMDDAITTDDTGTMTGFKSNYMRDLTRDGLHAPITTSKELNGVDAYVSNDRIY 300
 QY 249 LVQGTGVVYFLTKGTYLVSGYPKRLEKEVTPHGIILDSVDAFLCPSSRLHINAGR 308
 Db 301 LKNGDQVYIYKAGAHFTLLEGYPTLKEELN-----IEQVDAFAFCPGQRTAHIIQGRK 355
 QY 309 LWMULDKSAQAQATWTLPMPHEKVDGALCMKESLSGNSANGPGLYLHGPULYCYSDV 368
 Db 356 IYINILAAATPRBITLDAAPLPFGDIDAAFC-----SSDGKIKIFQGSNYHYHDS 403
 QY 369 EKLNAAKALPOPONVTS-LIGC 389
 Db 404 MLVMSRIAPILAKTISAMVGC 425

RESULT 4

Q90WF7 PRELIMINARY; PRT; 439 AA.
 AC Q90WF7; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE Warm-temperature-acclimation-related-65 kDa-protein precursor.
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kinoshita S., Itoi S., Watabe S.;
 RT "cDNA Cloning and Characterization of the Warm-Temperature-
 RT Acclimation-Related-65 kDa-Protein";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

```

Query Match      20.94; Score 456; DB 13; Length 261;
Best Local Similarity 36.24; Pred. No. 2.8e-30;
Matches 108; Conservative 46; Mismatches 92; Indels 52; Gaps 11;

Qy      96 PVDAAFROGNSVFLIKGDVWVYVPEPKK-KGYPLTLDPEFGIPSPDAAVECHRG 153
      ||||| : : : : :
Db      10 PCDAALR-WLERYCIGQGYFRFRPHSWEVLPCYPRDLNDYF--ICP----- 55

Qy      154 CQABGVLPFGCHGRNRGTGNGSTHGPCPEYMRCSPHVLVSLTSDNNGATYAEGETHYR 213
      ||||| : : : : :
Db      56 -----GRGR-----HGN-TSMWNAQDRCSGE-PFOALITSDSGHYAFRGSLFR 99

Qy      214 LPTSRDDGMSHPILAHQWPGSPSAVDAFAFMSBEKLYLQGTGVVFLR-KGGYTLVSGYRK 272
      ||||| : : : : :
Db      100 LDRSRDGMHAMPQAHSPGIGQGVDAFAFMDKRYTLQGSQVSLYVGGRGHQLVEGYPR 159

Qy      273 RLEKEVSTPHGIIILDSVDAFICPGSGRLIHMGRLMLDLKSGAQTWELFWPHKV 332
      ||||| : : : : :
Db      160 ALGELGVF-----KDAFAFTCGSALYVITGDRMQRVDTLTKSPRHADPEQPLPYDGV 213

Qy      333 DGALEKMSLGRPSCSANGPGVYLHGPRLNYCSDVDEKLAARLPPQPNV-TSLDLC 389
      ||||| : : : : :
Db      214 DGAMC-----TADGIYLLRGRDXYRHMDVALLAHPADPPSLAVDLFHC 259

RESULT 6
Q090310 PRELIMINARY; PRT; 445 AA.
AC Q090310;
DT 01-NOV-1996 (TREMBLrel, 01, Created)
DT 01-NOV-1996 (TREMBLrel, 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DS WARM temperature acclimation-related 65-kDa protein precursor.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
NC NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle, and Hepatopancreas;
RX MEDLINE=95340486; PubMed=7615502;
RA Kikuchi K., Yamashita M., Watabe S., Alda K.;
RT "The warm temperature acclimation-related 65-kDa protein, Wap65, in
   goldfish and its gene expression.";
RL J. Biol. Chem. 270:17087-17092(1995).
RN [2]
RP SEQUENCE OF 1-149 FROM N.A.
RC TISSUE=Hepatopancreas;
RA Kikuchi K., Watabe S., Alda K.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D50437; BAA08928.1; -.
DR EMBL; D85428; BAA19835.1; -.
DR HSSP; P20058; IHXN.
DR InterPro; IPR000585; Hemopexin.
DR Pfam; PF00045; hemopexin; 5.
DR SMART; SM00120; HX; 4.
KW Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 445 WARM TEMPERATURE ACCLIMATION-RELATED 65-
FT FT KDA PROTEIN, WAP65.
FT FT
SEQUENCE 445 AA; 50769 MW; 79C32EAB63BA82C6 CRC64;

Query Match      20.54; Score 447.5; DB 13; Length 445;
Best Local Similarity 28.64; Pred. No. 2.9e-29;
Matches 132; Conservative 68; Mismatches 153; Indels 109; Gaps 16;

Qy      15 SLGCSLAITAPRLPTSAHGVAERETPDVDV-----TERCSGSGFDTATLDNGTML 68
      ||||| : : : : :
Db      8 TLCLASLISLAASBHHKEDHYQDDEPRGHGHEHLHGANDLRCC-GMEFDALVAVBEGIPY 66

Qy      69 PFKGEFVWKSHKMDRELISRMKNP-----SPVDAAPRQ-----GHNSVFLIKG 113

```

```

Db 67 FPKDHLFGKHDOALSNB---TPRELDEHHHLGHVDAAFRMSSEDSPAHHDHQFFFLD 123
Qy 114 DKWVWVPEKKEKGYRLDDEPGISPLDAAVECHRGEOAGVLPFG-- 164
Db 124 TKVFSYKHLKEXDYPDISELFPGLPDHLDAAVECPDPCANDTIIFFEGDEIYHDMK 183
Qy 165 -----HGR-----NGTGHGNSTHGPEX-MRCS 188
Db 184 TKVDEKPKSMRSTGAFRMDHYGPHGHPKPRPIGEOGKPKXTRODTPMC-P 242
Qy 189 H-----LVLSALTSNKGATYAFSGTHYRLDTSRDGMHSPDAHQP 231
Db 243 HFGQKTDEHIEREQSRVHLDAITSDDSVAFRGNH--LSITGDKFHSDTIESAFK 300
Qy 232 QGSAVDAAFSMEKLYLVGTOYVFLTKGYTLVSGYPRLEKEVTHGIIILDSVDA 291
Db 301 ELHSEVDAVFSYEGHLMWIKNEVYKVGEPHNLHGYPRLKEVLG-----LEGVDA 355
Qy 292 AFICPGSSRLHIMAGRLMMLDLKSGAQTWELPMPH-EKVDGALCMKSLGPNCSAN 350
Db 356 AFVCAADHIAHVAGVTVVDLKAATPRVYKESINHLKIDVAMC-----GPK----- 405
Qy 351 GPELYLHGPNLYCYSDVEKLNAAKALPQPNVT-SLLGCTH 391
Db 406 --GVTAVIGHNYGFGSPMIMMAKIMPEQHRVSGLFGCDH 445

RESULT 7
ID Q8JIP9 PRELIMINARY; PRT; 430 AA.
AC Q8JIP9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Warm-temperature-acclimation-related-65 KDa-protein.
GN Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adriantichthyidae; Oryziatidae; Oryziatidae.
OC NCBI_Taxid=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HNI;
RA Hirayama M., Kobiyama A., Kinoshita S., Watabe S.;
RT "Primary structural analysis of medaka Wap65 and Wap65-like protein
RT and their gene expression in association with temperature acclimation
RT during ontogeny."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB075198; BAB97303.1;
DR InterPro: IPR000585; Hemopexin.
DR Pfam: PF00045; hemopexin; 2.
DR SMART: SM00120; HX; 4.
SQ SEQUENCE 430 AA; 49847 MW; A8BFC56B33AAEF15 CRC64;

Query Match 20.4%; Score 445.5; DB 13; Length 430;
Best Local Similarity 28.1%; Pred. No. 4e-29;
Matches 126; Conservative 67; Mismatches 155; Indels 101; Gaps 14;

Qy 16 LCWSLAATPLPPTSAHGNVAGETKPPDVTTERCSGDSWSPDATTLDNGTMLFFKGEFV 75
Db 10 LCLALVLA---ADHHEHRRKA-----VRDCK-GIEMDAVAVNEBGIPIYFKDHL 58
Qy 76 WKSHKMDRELISERKNP-----SPVDAFR-----QGHNSVFLIKGDKVWVPEK 123
Db 59 FKGHGQALSNKSFALDDHHLGHVDAAFRMATYEDLDNHDMPFLDNKVAAYQHK 118
Qy 124 KEKGYRLDDEPGISPLDAAVECHRGEOAGVLPFG-- 164
Db 119 LEAGYPAISEVFPGLPDHLDAAVECPDPCANDTIIFFEGDEIYHDMK 178
Qy 165 -----HGR-----NGTGHGNSTHGPEX-MRCS----- 187

```

```

Db 179 SMPNCTSAFRMEHYCFGHKSKEDPKTGEVRGKYPKDAKFFMRCSFEDNDHEER 238
Qy 188 ---PHLVLSALTSNKGATYAFSGTHYRLDTSRDGMHSPDAHQPQBSAVDAAFSWE 244
Db 239 ERCSYHVLDAITSDDSVAFRGNH--LSITGDKFHSDTIESAFKELHSEVDAVFSYN 298
Qy 245 EKLIVGQNVYVFLTKGYTLVSGYPRLEKEVTHGIIILDSVDAAFICPGSSRLHIM 304
Db 299 SHLYMKDQDLFYRGEPRHLAGIPKRYOALG-----IKPIDAAPVCOBRIAHII 353
Qy 305 AGRRLMMLDLKSGAQTWELPMPH-HEKVDGALCMKSLGPNCSANGELYLHGNLY 363
Db 354 KDRHMYDVMSAFPRATNRPRISILKVDGAMC-----GPK-----GVKVRGNHY 401
Qy 364 CYSDEVKLNAAKALPQPNVT-SLLGCTH 391
Db 402 HFESPKTFVAARALPQHRISLELFGCDH 430

RESULT 8
ID P79825 PRELIMINARY; PRT; 446 AA.
AC P79825;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hemopexin-like protein (fragment). (Salmo gairdneri).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OC NCBI_Taxid=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97142135; PubMed=8988369;
RA Mot S., Duval J., Le Goff P.;
RT "Molecular cloning of a hemopexin-like cDNA from rainbow trout
RT liver."
RL DNA Seq. 6:311-318(1996).
DR EMBL: Z68112; CA92147.1;
DR HSSP: P20058; IHNX.
DR InterPro: IPR000585; Hemopexin.
DR Pfam: PF00045; hemopexin; 4.
DR SMART: SM00120; HX; 4.
DR NON TER 1
SQ SEQUENCE 446 AA; 50454 MW; 13A8BA4E42B37227 CRC64;

Query Match 20.2%; Score 442; DB 13; Length 446;
Best Local Similarity 27.8%; Pred. No. 8.3e-29;
Matches 126; Conservative 64; Mismatches 168; Indels 96; Gaps 13;

Qy 16 LCWSLAATPLPPTSAHGNVAGETKPPD-----VTERCSGDSWSPDATTLDNGTMLFF 70
Db 11 LCLVLAISHAHNHAHQGGDEBHGHDHGBGLLDRC-QGIEMDAVAVTEBGIPIYF 69
Qy 71 KGFVWKSMDRELISERKNP-----SPVDAFR-----QGHNSVFLIKGDKVW 118
Db 70 KGHVFPKFGKALSNKSFALDDHHLGHVDAAFRMATYEDLDNHDMPFLDNKVAAYQHK 129
Qy 119 YPEKKEKGYRLDDEPGISPLDAAVECHRGEOAGVLPFG-- 164
Db 130 YKHQLETFPPKQISEVFPGLPDHLDAAVECPDPCANDTIIFFEGDEIYHDMK 189
Qy 165 -----HGR-----NGTGHGNSTHGPEX-MRCS----- 187
Db 190 EKKFEGMPCNTSAFRMEHYCFGHGQFSGKFDKTBGVGRYKEARDFMKSGKGD 249
Qy 188 -----PHLVLSALTSNKGATYAFSGTHYRLDTSRDGMHSPDAHQPQBSAVDA 239
Db 250 DHIERCSRVHLDAITSDDSVAFRGNH--LSITGDKFHSDTIESAFKELHSEVDA 309

```

QY 240 AFSEMEKLYVGTGVYFLTKGGYTLVSGYPRKLEKEVGTGPIILDSVDAAFICPGSS 299
 DB 310 TFSYENHLVYVKKODKYIYKVGSHHLDSPKRLKEVLG-----VEGPIIDAFVQDDH 364
 QY 300 RLHIMAGRRLMIDLKSGAQTWTELPW-HEKYDGLCMESLGPNSCANGPGLYLH 358
 DB 365 IAHVHKQYVYDVLKASPPVAVESFPLFNKVDAMC-----GPE-----GVGLFK 412
 QY 359 GPNLYCSDVEKLNAAKALPQ-ONVTSLLGCTH 391
 DB 413 GNHYFHFQSVKVMIMAKALPEEHKTALELFGCDH 446

RESULT 9

Q9DFN1 PRELIMINARY; PRT; 211 AA.
 AC Q9DFN1; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hemopexin-like protein (Fragment).
 OS Gilllichthys mirabilis (Long-jawed mudsucker).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Gobioidi;
 OC Gobiidae; Gilllichthys.
 NC NCB1_TaxID=8222;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=21117151; PubMed=11172064;
 RA Gracey A.Y., Troll J.V., Somero G.N.;

RT "Hypoxia-induced gene expression profiling in the euryoxic fish
 RT Gilllichthys mirabilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).
 DR EMBL; AF266204; AAG13324.1; -.
 DR HSSP; P20058; IHN.
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 4.
 DR SMART; SM00120; HX; 2.
 FT NON TER 1
 SQ SEQUENCE 211 AA; 24104 MW; EC149A2AC07B13CC CRC64;

Query Match 12.4%; Score 271; DB 13; Length 211;
 Best Local Similarity 33.0%; Pred. No. 7.2e-15;
 Matches 69; Conservative 37; Mismatches 83; Indels 20; Gaps 6;

QY 185 RCGPMLVLSALTSPDNHGAATAFSGTHYMRLDTSRDGMHWPIDAHQWPGSAVDAAFSWE 244
 DB 21 KCS-RVHLDDITSDNAGMNTAFRPHHLEKDEGNDITVDNIEAFELHSDVDAYFSYD 79
 QY 245 EKLTVGTGVYVFLTKGGYTLVSGYPRKLEKEVGTGPIILDSVDAAFICPGSSRLHIM 304
 DB 80 DHLYMIDKHVHYKVAEPHTHLEGYKTYEELG-----IDGVDAAPCEBDHINHF 134
 QY 305 AGRLIMLIDLKSGAQTWTELPW-HEKYDGLCMESLGPNSCANGPGLYLHGNLY 363
 DB 135 QGKNIFFVNDKVSREPTEPRLALFKKVDAMC-----GPG-----GLKVVGNHFF 182
 QY 364 CYSDEKLNAAKALPQ-ONVTSLLGCTH 391
 DB 183 RFDVYMLVAGRALPEQHRVSTELFGCDH 211

RESULT 10

Q9DFN1 PRELIMINARY; PRT; 116 AA.
 AC Q9DFN1; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hemopexin-like protein variant 1 (Fragment).
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Proteacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NC NCB1_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;

RT "Immune-relevant (including acute phase) genes identified in the
 RT livers of rainbow trout, Oncorhynchus mykiss, by means of suppression
 RT subtractive hybridization.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF281339; AAG30014.1; -.
 DR HSSP; P45452; IPEX.
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 1.
 DR SMART; SM00120; HX; 1.

FT NON TER 1
 SQ SEQUENCE 116 AA; 13379 MW; 4543D625B162B3BC CRC64;

Query Match 10.3%; Score 224; DB 13; Length 116;
 Best Local Similarity 40.7%; Pred. No. 3e-11;
 Matches 46; Conservative 20; Mismatches 35; Indels 12; Gaps 2;

QY 64 NGTMLPFKGFYVKKSHKMDRELISERKNPPS-----PVDAPR-----QGNHGVFLI 111
 DB 1 NGHTFFKGDHMLNGKGPAAQVSSAFPKELDYNHGHVDAAPRMNKKERPEGDHITFR 60
 QY 112 KGDVWVYPRPEKKKGYPKLDDPPGIPSPDAAVECHGECAQVLEFQ 164
 DB 61 LDKKVSYYNHSLSEGYPPDIOQDPGVPVSHVDAVECPKGCNSDVSLEFKG 113

RESULT 11

QY 012945 PRELIMINARY; PRT; 453 AA.
 AC 012945;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Vitronecclin.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCB1_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn; TISSUE=Retina;
 RX MEDLINE=98090066; PubMed=9362471;
 RA Martinez-Morales J.R., Barbas J.A., Marti E., Bovolenta P., Edgar D.,
 RA Rodriguez-Tebar A.;

RT "Vitronecclin is expressed in the ventral region of the neural tube and
 RT promotes the differentiation of motor neurons.";
 RL Development 124:5139-5147(1997).
 DR EMBL; Y11030; CA471914.1; -.
 DR HSSP; P45452; IPEX.
 DR InterPro; IPR000585; Hemopexin.
 DR InterPro; IPR001212; Somatomedin_B.
 DR Pfam; PF00045; hemopexin; 4.
 DR Pfam; PF01033; Somatomedin B; 1.
 DR PRINTS; PRO00022; SOMATOMEDINB.
 DR SMART; SM00120; HX; 4.
 DR SMART; SM00201; SO; 1.
 DR PROSITE; PS00024; HEMOPEXIN; 2.
 DR PROSITE; PS00524; SOMATOMEDIN B; 1.
 SQ SEQUENCE 453 AA; 51636 MW; 82B3BDC2B9397C4 CRC64;

Query Match 9.2%; Score 200.5; DB 13; Length 453;
 Best Local Similarity 25.8%; Pred. No. 1.6e-08;
 Matches 91; Conservative 42; Mismatches 103; Indels 117; Gaps 20;

QY 26 LPPTSAHGNVAEGRTK---DPVTERCSGSGFADATYLLDDNCTMLPFKGEFVWKS HKMD 82

```

Db      127 LHPITTTTSDTRNPSLDDEBELCSRK-PFNAFTDLNKGSIYAFRGKYFELDKSS 185
Qy      83 -----RELISERKNFSPVDAFR-----OGHNSVFLIKGDKVWVPEKKEKGYPKLQ 133
Db      186 VRBGYPKLISDW-GIGSPIDALFTRINCOG-KTYLFKSGQYRFPDDGALDIPYDIDIS 242
Qy      134 DEFGIIPSLDAVECHRGEOAGVLFFQGHGRNGTGHNSTHHGEVYRCSPLVLS 193
Db      243 EGFGIIPNDIDA-----FALPAH-----SYHGNE-----267
Qy      194 ALTSDNNGATYAFEGTTHWRLDTSRDGHSWPIAHQMPQ-----GSAV-DAAF-----241
Db      268 -----RVYFFKGYTW-----SYDFAHQPTQAECEKSSPSYTFNHAFAWRD 309
Qy      242 SWEKLYVVOGTQYVFLTKGTYLVSGYPRKLEKE--VGTPHGIILSDVDAAFICPSS 299
Db      310 SWED-----IFLSFGSRMVGASSORLISRDWGVNQ--LDAMAGRIYSSR 356
Qy      300 RLHIMAGR-----LM-WDLKSGAQTWELPMPHEKVDGALC 337
Db      357 QPRRRSSRRHRKRYNRHTNLGLMSWLN--SDSESTDSDW-----LSGSQC 403

```

RESULT 12

```

Q81WC3 ID Q81WC3 PRELIMINARY; PRT; 634 AA.
AC Q81WC3;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strauberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040507; AA040507.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 634 AA; 70086 MW; 0D4A7863BAACAF CRC64;

```

Query Match 9.1%; Score 199; DB 4; Length 634;

Best Local Similarity 25.0%; Pred. No. 3.2e-08; Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

```

Qy      5 LGAPVALGL-----WSLCWSLAITPL-----PPTSAHGNVAEGETKPDPPVTERCS 51
Db      305 VGDPRLRGLEVEBKVRWQLYGVRESVPTAQPEFPLBPBPNNRSAPPRKDVPIRCS 364
Qy      52 DGSFPAATLDDNGTMLPFKGEFVWKSMDRELISER-----WKNFP--SPVDAAF 101
Db      365 T--HFDVA-QIRGEAFFKGYFWRLTR--DRHLVSLQAPQMRFWGLPLHLSDVAVY 420
Qy      102 -RGHNSVFLIKGDKVWVPEKKEKGYPKLQDEFGIIPSLDAVECHRGEOAGVL 160
Db      421 ERKSDHKTIVFKGDRVWFVFDNNVEBGYPVSD-FSLPFGIDAA-----465
Qy      161 FFOGHGRNGTGHNSTHHGEVYRCSPLVLSALTSDNNGATYAFEGTTHWRLDT--SRD 219
Db      466 FSAHADR-----TYFFQDLWRYDHTRH 491
Qy      220 GHSWPI-AHQWPGSAVDAAFSWEK-KLYVQGTQVYVFLTKGTYLVSGYPRKLEKE 277
Db      492 MDGYPKQSPILWGVSTLDDAMKMSDGAIFYFRG-QEYWKVLDELGLVAPGYFGSTARD 550
Qy      278 -----VGTPIGIIILSDVDA 292
Db      551 WLVCGSDQADGSVAAGVDAA 570

```

RESULT 13

```

Q99745 ID Q99745 PRELIMINARY; PRT; 240 AA.
AC Q99745;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Metalloproteinase produced by human alveolar macrophages."
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94043200; PubMed=8226919;
RA Shapiro S.D., Kobayashi D.K., Ley T.J.;
RT "Cloning and characterization of a unique elastolytic
RT metalloproteinase produced by human alveolar macrophages."
RL J. Biol. Chem. 268:23824-23829 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97187267; PubMed=9034720;
RA Borden P., Heller R.A.;
RT "Transcriptional control of matrix metalloproteinases and the tissue
RT inhibitors of matrix metalloproteinases."
RL Crit. Rev. Eukaryot. Gene Expr. 7:159-178 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Lin D., Duncan M., Allen E., Araujo R., Aparicio A., Chai A.,
RA Chung E., Davis K., Federspiel N., Hyman R., Kaiman S., Kemp C.,
RA Kurd O., Lashari D., Lew H., Namath A., Oefner P., Roberts D.,
RA Heller R., Davis R.W.;
RT "Three Matrix Metalloproteinases on 81kb of Pl insert."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78045; AAB36943.1; -.
DR HSSP; P39900; IJIZ.
DR Interpro; IPR000585; Hemopexin.
DR Pfam; PF00045; hemopexin; 4.
DR SMART; SM00120; HX; 4.
DR PROSITE; PS00024; HEMOPEXIN; 1.
FT NON_TER 1
SQ SEQUENCE 240 AA; 28406 MW; 3E620FAF17EB7CCC CRC64;

```

Query Match 9.0%; Score 196.5; DB 4; Length 240;

Best Local Similarity 26.9%; Pred. No. 1.5e-08; Matches 65; Conservative 22; Mismatches 92; Indels 63; Gaps 10;

```

Qy      30 SAHGNVAEGETKPD--VTERCSDSGSPDATLDDNGTMLPFKGEFVW-----KSHKMDR 83
Db      30 SLVGDPKRNQRLPNPNKSPALCDPMLSDPAVTVGN-KIFFKDNFPMVKVSEPKTSV 88
Qy      84 ELISERKNFSPVDAFR-OGHNSVFLIKGDKVWVPEKKEKGYPKLQD-EFGIIPS 141
Db      89 NLISLSMPTLPSSGIEAAYEIKARNOVFLFKDDQYWLISMLRPPNPYKSIHSPGFNPFV 148
Qy      142 PLDAVECHRGEOAGVLFFQGHGRNGTGHNSTHHGEVYRCSPLVLSALTSDNHG 201
Db      149 KIDAAV-----FNRFR-----161
Qy      202 ATYAFSGTTHWRLDTSRD-----GHSWPIAHQWPGSAVDAAFSWEK-LYVQGTQVY 256
Db      162 -TYFVYDQYWRDERQWMDPEYPLITKNGGICPK-IDAVFYKKNKYYVFGQSNQF 219
Qy      257 VF 258
Db      220 EY 221

```

RESULT 14

```

Q9TV75 ID Q9TV75 PRELIMINARY; PRT; 464 AA.

```



```

AC Q9TV75;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Matrix metalloproteinase-12 precursor.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Japan White; TISSUE=long bones;
RX MEDLINE=98424349; PubMed=9753427;
RA Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,
RA Kawashima H.;
RT "Large scale isolation of osteoclast-specific genes by an improved
RT method involving the preparation of a subtracted cDNA library.";
RL Gene Cells 3:459-475(1998).
DR HSSP; P39900; 1J1Z.
DR EMBL; AB006779; BAA75922.1; -.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Matrixin.
DR InterPro; IPR006026; Nzn_MTPeptide.
DR InterPro; IPR006025; Zn_MTPeptide.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10_N; 1.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HK; 4.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR HydroLase; Metal-binding; Metalloprotease; Protease; Signal; Zinc.
KW SIGNAL
FT 1
FT 17
SQ SEQUENCE 464 AA; 53585 MW; B9A81A391E24317E CRC64;

Query Match 9.0%; Score 196; DB 6; Length 464;
Best Local Similarity 27.4%; Pred. No. 3.8e-08;
Matches 64; Conservative 19; Mismatches 91; Indels 60; Gaps 8;

QY 30 SAHGNVAEGETKDPDUTE--RCSGMSFPAATLDDNGTMLFKGGEFVW---KSHKMR 83
DB 255 SLXGEGHQPMKPNPEPTACDHLKFDAYTVGN-KLFFKDSFFWKKIPKSTTSV 313
QY 84 ELISERWKNPSPVDAAFRQ--HNSVFLIKGRVWVYPPPEKKEGYPKLQD--EPPGIPS 141
DB 314 RLISLIMPTLPSGIRAAVEIGDRHQVFLFKGDKFMLISHLRLQPNYPKSIHSLGFPDPVK 373
QY 142 PLDAAVECHRGECQAGVLFPGHGHNGTGHGNSTHHGPEYMRCSPHLVLSALTSDNHG 201
DB 374 KIDAAV-----FNPSL-----R 385
QY 202 ATYAFGSTHYWRDLTSS---DGWHSWPIAHQWPGSPSAVDAAPSWEEKLYLVQG 252
DB 386 KTYFVDNLVWRDERREVMAGYPKLITGHFPGIGPKIDAVYFYQRYTYFFQG 439

```

```

RA Sorensen K.C., Kitchell B.E.;
RT "Comparison of feline stromelysin-1 (MMP3) cDNA with other known MMP3
RT cDNA sequences.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033627; AAK64605.1; -.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Matrixin.
DR InterPro; IPR006026; Nzn_MTPeptide.
DR InterPro; IPR006025; Zn_MTPeptide.
DR Pfam; PF00045; hemopexin; 3.
DR Pfam; PF00413; Peptidase_M10_N; 1.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HK; 3.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 393 AA; 44093 MW; 148C78BAEBDA95 CRC64;

Query Match 8.7%; Score 189.5; DB 6; Length 393;
Best Local Similarity 26.4%; Pred. No. 1.1e-07;
Matches 63; Conservative 23; Mismatches 94; Indels 59; Gaps 8;

QY 25 PLPPTSAHGNVAEGT-KDPDUTERCSDGMSFPAATLDDNGTMLFKGGEFVW---KSHKMR 83
DB 201 PPPMSSPDGPAGPTTSVPEPETPATCDPALSFDVVS-SLRGELLFPKDRHLMRKSLRTR 259
QY 84 E---LISERWKNPSPVDAAFRQ--HNSVFLIKGRVWVYPPPEKKEGYPKLQD--EPP 137
DB 260 EPGFILMCSFWSLSLSDGDAVEBTSKQIVLFPKNGFWAMRGTEVQAGYPKGHTLGRF 319
QY 138 GLPSPLDAAVECHRGECQAGVLFPGHGHNGTGHGNSTHHGPEYMRCSPHLVLSALTS 197
DB 320 PYVKKIDAAV-----FD 331
QY 198 DNHGATYAFSGTHYWRDLTSSRDGH-SWP--IAHQWPGSPSAVDAAPSWEEKLYLVQG 253
DB 332 KKKKTYFVGVGKWRFEKRSMEBPGPKQIAEDPGVDISVDAFAEAFGYFFPENG 390

```

Search completed: December 16, 2003, 06:52:11
Job time : 179 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:37:44 ; Search time 64 Seconds
(without alignments)
969.720 Million cell updates/sec

Title: US-09-900-448-2
Perfect score: 2185
Sequence: 1-MARVLGAPVALGLMGLCWSL.....NAKALPQPNTSLGCTH 391

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2139.5	97.9	462	22	AAAM23933
2	456	20.9	83	21	AAG00304
3	258	11.8	87	23	ABP31577
4	204	9.3	587	21	AAAY90502
5	204	9.3	587	21	AAV88405
6	202.5	9.3	470	22	AAE10419
7	202.5	9.3	470	22	AAAB84614
8	202.5	9.3	470	22	AAAB74595
9	202.5	9.3	470	22	AAAB49982

10	202.5	9.3	470	23	AAU91061	Human head and nec
11	202.5	9.3	470	24	ABP96800	Human COPD related
12	202.5	9.3	470	24	ABR48203	Human bladder canc
13	202.5	9.3	470	24	ABU56421	Lung cancer-associ
14	202.5	9.3	470	24	ABU56663	Lung cancer-associ
15	202.5	9.3	473	21	AAAB3772	Human cancer-associ
16	199	9.1	519	22	AAE10426	Human matrix meta
17	199	9.1	519	22	AAE10426	Amino acid sequenc
18	199	9.1	605	21	AAAY90503	Human transmembran
19	199	9.1	605	24	AAO19713	MT4-WMP catalytic
20	199	9.1	606	21	AAV88404	Human MT4-WMP (2)
21	196	9.0	464	21	AAW52135	Rabbit matrix meta
22	195.5	8.9	477	23	AAU91062	Human head and nec
23	192.5	8.8	476	23	ABU03460	Angiogenesis-asso
24	189.5	8.7	475	23	ABU05574	Breast cancer-asso
25	189.5	8.7	475	23	ABU56595	Lung cancer-associ
26	189.5	8.7	476	22	AAE10421	Human matrix meta
27	189.5	8.7	476	22	AAE10421	Amino acid sequenc
28	189.5	8.7	476	24	ABR48147	Human bladder canc
29	189.5	8.7	476	24	ABP54455	Matrix metalloprot
30	183.5	8.4	444	22	AAE10416	Human WMP-galt pol
31	183.5	8.4	467	22	AAE10416	Human matrix meta
32	183.5	8.4	467	22	AAE10416	Amino acid sequenc
33	183.5	8.4	467	22	AAE10416	Human neutrophil c
34	181.5	8.3	469	22	AAE10415	Human matrix meta
35	181.5	8.3	469	22	AAE10415	Amino acid sequenc
36	181.5	8.3	469	24	ABR48148	Human bladder canc
37	181.5	8.3	469	24	ABU56596	Lung cancer-associ
38	181.5	8.3	469	24	ABU56597	Lung cancer-associ
39	181.5	8.3	469	24	ABU07454	Protein different
40	181.5	8.3	469	24	ABP54454	Matrix metalloprot
41	181.5	8.3	492	13	AAE24863	Sequence of pre-pr
42	181.5	8.3	495	22	AAE35509	Human colon cancer
43	180.5	8.3	457	10	AAE93628	Sequence of human
44	180.5	8.3	469	8	AAE70611	Sequence encoded b
45	179.5	8.2	479	23	ABP41861	Human ovarian anti

ALIGNMENTS

RESULT 1	
AAAM23933	standard; Protein: 462 AA.
ID	AAAM23933;
AC	AAAM23933;
XX	
DT	12-OCT-2001 (first entry)
DE	Human EST encoded protein SEQ ID NO: 1458.
XX	
KW	Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW	tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW	diagnostics; forensic test; gene mapping; genetic disorder;
KW	biodiversity; gene therapy; nutrition.
XX	
OS	Homo sapiens.
XX	
PN	W0200154477-A2.
XX	
PD	02-AUG-2001.
XX	
PR	25-JAN-2001; 2001WO-US02687.
XX	
PR	25-JUN-2000; 2000US-0491404.
XX	
PR	17-JUL-2000; 2000US-0617746.
XX	
PR	03-AUG-2000; 2000US-0631451.
XX	
PR	15-SEP-2000; 2000US-0663870.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX	
PI	Cao Y, Dimaic RA, Zhang J, Werhman T;

XX WPI; 2001-476164/51.
 DR N-PSDB; AAH98592.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 XX antibodies and research use -
 XX
 PS Claim 20; Page 1017-1018; 1275pp; English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.
 XX
 SQ Sequence 462 AA;
 XX
 Query Match 97.9%; Score 2139.5; DB 22; Length 462;
 Best Local Similarity 84.6%; Pred. No. 8.9e-193;
 Matches 391; Conservative 0; Mismatches 0; Indels 71; Gaps 1;
 QY 1 MARVLGAPVALGLMSLCMSLAIAIATPLPTSAHGNAEGETKPPDVTERTCSGMSFPAAT 60
 DB 1 MARVLGAPVALGLMSLCMSLAIAIATPLPTSAHGNAEGETKPPDVTERTCSGMSFPAAT 60
 QY 61 LDNGTMTLFFKGEFVWKSXKMDRELISERWKNFPPSDAARFQGHNSVFLIKDKRWVYP 120
 DB 61 LDNGTMTLFFKGEFVWKSXKMDRELISERWKNFPPSDAARFQGHNSVFLIKDKRWVYP 120
 QY 121 PEKKKGYPLLDDEPFGIPSPIDAARVCHRGCOAGVLPFQ----- 163
 DB 121 PEKKKGYPLLDDEPFGIPSPIDAARVCHRGCOAGVLPFQDRNEMFDLATGTMKER 180
 QY 164 -----GHGHRN 169
 DB 181 SWPAVNCCSALMLWGLGRYCFQGNQFLRPDPVAVGEVPRPRVROVDFMPCPGRGHNRN 240
 QY 170 GTCHGNSHTHGPYMRCSPLVLSALTSDNHGATVAFSGTHYWRDLTSDRDMHSWPIAQ 229
 DB 241 GTCHGNSHTHGPYMRCSPLVLSALTSDNHGATVAFSGTHYWRDLTSDRDMHSWPIAQ 300
 QY 230 WPGSPAVDAAFSMEKLYVVOGTQVYVFLTKGYTLVSGYPRLEKEVGTPIGILLDSV 289
 DB 301 WPGSPAVDAAFSMEKLYVVOGTQVYVFLTKGYTLVSGYPRLEKEVGTPIGILLDSV 360
 QY 290 DAAFIQPGSSRLHIMAGRRLMWLDLKSQAQATWTELPWPEKVDGALCMEKSLGPNSSCA 349
 DB 361 DAAFIQPGSSRLHIMAGRRLMWLDLKSQAQATWTELPWPEKVDGALCMEKSLGPNSSCA 420
 QY 350 NGEGLYLHIGPNLYCYSDEVKLNAAKALPQPONVTSILGCTH 391
 DB 421 NGEGLYLHIGPNLYCYSDEVKLNAAKALPQPONVTSILGCTH 462
 XX
 RESULT 2
 AAG00304
 ID AAG00304 standard; Protein; 83 AA.
 XX
 AC AAG00304;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 4385.
 XX
 KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.

XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-020610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 DR N-PSDB; AAC00310.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 4385; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 83 AA;
 XX
 Query Match 20.9%; Score 456; DB 21; Length 83;
 Best Local Similarity 98.8%; Pred. No. 3.9e-35;
 Matches 82; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MARVLGAPVALGLMSLCMSLAIAIATPLPTSAHGNAEGETKPPDVTERTCSGMSFPAAT 60
 DB 1 MARVLGAPVALGLMSLCMSLAIAIATPLPTSAHGNAEGETKPPDVTERTCSGMSFPAAT 60
 QY 61 LDNGTMTLFFKGEFVWKSXKMDR 83
 DB 61 LDNGTMTLFFKGEFVWKSXKMDR 83
 XX
 RESULT 3
 ABP31577
 ID ABP31577 standard; Protein; 87 AA.
 XX
 AC ABP31577;
 XX
 DT 09-JUL-2002 (first entry)
 XX
 DE Human glycoprotein-like ORF550 protein, SEQ ID NO:1100.
 XX
 KM Human; ORF; open reading frame; ORF; drug screening; diagnosis;
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;
 KM immune modulation; haematopoiesis regulation; tissue growth;
 KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KM thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KM behaviour; cancer; proliferative disorder; neurological disorder;
 KM cardiovascular disease; immune system disorder; organ transplantation;
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KM hypothyroidism; cholesterol ester storage disease; infection; vulvareary;
 KM vasotropic; antiporiatic; antidiabetic; cytosolic; nocotropic;
 KM neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KM cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KM dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.
 XX MO200190366-A2.
 XX 29-NOV-2001.
 XX 24-MAY-2001; 2001WO-US17076.
 XX 24-MAY-2000; 2000US-206690P.
 XX (CURA-) CURAGEN CORP.
 XX Leach MD, Shimkets RA;
 XX WPI; 2002-106200/14.
 DR N-PSDB; ABN75603.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 XX Claim 10; Page 541-542; 2508pp; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, hematopoietic regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and anti-infective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 XX
 SQ Sequence 87 AA;
 Query Match 11.8%; Score 258; DB 23; Length 87;
 Best Local Similarity 86.5%; Pred. No. 2e-16; Indels 0; Gaps 0;
 Matches 45; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 114 DKVAVYPPKKKKYKYLQDEFGIPSPDLDAVECHRGECQAGVLPFGQH 165
 DB 5 DKVAVYPPKKKKYKYLQDEFGIPSPDLDAVECHRGECQAGVLPFGQH 56
 RESULT 4
 AA190502
 ID AA190502 standard; Protein; 587 AA.

XX
 AC AA190502;
 XX
 DT 15-AUG-2000 (first entry)
 XX
 XX Murine transmembrane matrix metalloprotease MT4-MMP(2).
 XX
 XX Matrix metalloprotease; MT4-MMP(2); transmembrane; mouse; antibody;
 XX drug screening; diagnosis; arthritis deformans; rheumatoid arthritis;
 XX asthma; autoimmune disease; atopic dermatitis; skin disorder;
 XX ischaemic disorder; arteriosclerosis; cancer; tissue damage;
 XX inflammatory disorder; neurological disorder.
 XX
 OS Mus sp.
 XX
 XX MO200018805-A1.
 XX
 XX 06-APR-2000.
 XX
 XX 29-SEP-1999; 99WO-JP05350.
 XX
 XX 29-SEP-1998; 98JP-0291501.
 XX 29-SEP-1998; 98JP-0291503.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 XX Hanai N, Furuya A;
 XX WPI; 2000-293115/25.
 DR N-PSDB; AAA14339.
 XX
 PT Antibody immunospecific for transmembrane matrix metalloprotease
 PT polypeptide, useful in screening of inhibitors and activators as drugs
 PT for, and for diagnosis of e.g. arthritis deformans and asthma -
 XX
 XX Claim 1; Page 86-90; 152pp; Japanese.
 XX
 CC This sequence represents a novel murine transmembrane matrix
 CC metalloprotease MT4-MMP(2). MT4-MMP(2) has physiological activity
 CC different to that of MT4-MMP. The invention relates to antibodies
 CC against both murine and human MT4-MMP(2) (AA190502 and AA190503) and to
 CC antibodies against a novel murine and human transmembrane matrix
 CC metalloprotease MT5-MMP (AA190504 and AA190505). The antibodies are
 CC useful for the diagnosis and screening of inhibitors and activators
 CC useful for treating or preventing a wide variety of medical conditions,
 CC such as arthritis deformans, rheumatoid arthritis, asthma, autoimmune
 CC diseases and atopic dermatitis. The anti-MT5-MMP antibodies may
 CC additionally be used to diagnose and screen therapeutic or prophylactic
 CC agents for conditions affecting the brain, such as cerebral stroke and
 CC Alzheimer's disease. The antibodies of the invention may also be used for
 CC the prevention, diagnosis and treatment of psoriasis, contact
 CC dermatitis, hair loss, ischaemic diseases, immune reaction accompanying
 CC organ transplant, hepatitis, nephritis, pancreatitis, arteriosclerosis,
 CC leukaemia, malignant tumours, wounds, corneal ulcers, tissue damage or
 CC inflammation accompanying leukocytic infiltration, dementia, multiple
 CC sclerosis, Parkinson's disease or brain tumour.
 XX
 SQ Sequence 587 AA;
 Query Match 9.3%; Score 204; DB 21; Length 587;
 Best Local Similarity 24.4%; Pred. No. 3.9e-10;
 Matches 103; Conservative 34; Mismatches 127; Indels 158; Gaps 25;
 QY 5 LGAPVALGL-----WSLCWGLAATATP-----PPISAGNVAEGSTKDDPV 46
 DB 273 VGPVVRVGLPEYEDRVVRWQLYGVRESVSPQAQDTPBPBPPLPBPBPNNRSTPPQKDV 332
 QY 47 TERCSQMSDPDATTLNDNGTMLPFKGEFVKSHKMPRELISER-----WKXFP---SP 96
 DB 333 PNRCTA--HFDAVA-QIRGEAPFFKGYFWRLLTR-DHLVSLDPAOMGRWRGLPLHDS 388
 QY 97 VDAAF-ROGNSVFLIKGDKVWVYPPKKKKYKYLQDEFGIPSPDLDAVECHRGECQ 155

Db 389 VDAVERTSDHKIVFPGDRYVWFKDNVVEGYRPRVSD-FSLPGGIDA----- 437
 QY 156 AEGVLFPGQGHNGHNGSTHGGPRVWCSPHLVLSALTSNKGATVAFSGTHWRD 215
 Db 438 ---VFSWANDR-----TYFFKQQLWRYD 459
 QY 216 --TSR--DGWHS-WPIAHQWPGPSAVDAFAFSWEE-KLYLVQGTQVYVFTKGGYTLVSG 269
 Db 460 DHTRRMDPGYPAGGPL---WRGVPSMLDDMRMSDGSATFFRG-QEYWKVLDSLELAAPG 515
 QY 270 YPRLEKE---VGTPHGILLSDVDAFICPGSSRLHIMAGRRLWLDLKSQAQATWTELP 326
 Db 516 YPGSTARMDLVGCEP-----LADAEVGPQPQ-----GR-----SGAQ----- 548
 QY 327 WPHKXVDG-ALCMKSLGPRNSCANGPGLYLHGNPNLYCYSDVEKLNAAKALPQPNVTS 385
 Db 549 ---DGLAVCS-----CTSDAHL---ALPSLLLTLP 573
 QY 386 LL 387
 Db 574 LL 575
 RESULT 5
 AAY88405
 ID AAY88405 standard; Protein; 587 AA.
 AC AAY88405;
 XX 28-JUL-2000 (first entry)
 DE Mouse MT4-MMP (2) matrix metalloproteinase amino acid sequence.
 XX
 KM Matrix metalloproteinase; MT4-MMP; treatment; arthritis deformans;
 KM Chronic Rheumatoid arthritis; asthma; autoimmune disease; hepatitis;
 KM atrophic dermatitis; psoriasis; contact dermatitis; hair loss; nephritis;
 KM ischaemic disease; pancreatitis; arteriosclerosis; leukaemia; tumour;
 KM wound; corneal ulcer; tissue damage; inflammation; cerebral apoplexia;
 KM Alzheimer's disease; dementia; multiple sclerosis; Parkinson's disease;
 KM gene therapy.
 KM
 OS Mus sp.
 XX
 PN W0200018900-A1.
 XX
 PD 06-APR-2000.
 XX
 PF 29-SEP-1999; 99WO-JP05349.
 XX
 PR 29-SEP-1998; 98JP-0276258.
 PR 29-SEP-1998; 98JP-0291505.
 XX
 PA (SEIKI) SEIKI M.
 XX
 PI Seiki M;
 XX
 DR WPI: 2000-293141/25.
 DR N-PSDB; AAA13378.
 XX
 PT Physiologically-active transmembrane matrix metalloproteinase
 PT polypeptide, useful in screening inhibitors and activators for treating
 PT e.g. arthritis deformans, asthma and cancers -
 XX
 PS Claim 11; Page 59-66; 113pp; Japanese.
 XX
 CC This sequence represents the murine matrix metalloproteinase (MT4-MMP 2)
 CC amino acid sequence. The invention relates to a mammalian transmembrane
 CC MMP polypeptide, its variants and the DNA sequences encoding them. Also
 CC included in the invention is a vector comprising the MMP DNA molecules,
 CC and cells transformed using the vector. The protein, DNA and antibodies
 CC directed against the polypeptide are useful in screening inhibitors and
 CC activators for use in treating arthritis deformans, chronic rheumatoid
 CC arthritis, asthma, autoimmune diseases, atrophic dermatitis, psoriasis,

CC contact dermatitis, hair loss, ischaemic diseases, immune reaction
 CC accompanying organ transplants, hepatitis, nephritis, pancreatitis,
 CC arteriosclerosis, leukaemia, malignant tumours, wounds, corneal ulcers,
 CC tissue damage or inflammation accompanying leukocytic infiltration,
 CC together with brain disorders during cerebral apoplexia, Alzheimer's
 CC disease, dementia, multiple sclerosis, Parkinson's disease or brain
 CC tumours. The DNA can also be incorporated into a suitable vector for use
 CC in gene therapy.
 XX
 SO Sequence 587 AA;
 Query Match 9.3%; Score 204; DB 21; Length 587;
 Best Local Similarity 24.4%; Pred. No. 3.9e-10;
 Matches 103; Conservative 34; Mismatches 127; Indels 158; Gaps 25;
 QY 5 LGAPVALGL-----WELCMSLATATPL-----PPRSAGNVAAEGTKPDPU 46
 Db 273 VGDPRVYGLPEYEDRVWOLYGVRESVSPQAOLDPEPEEPPLPPEPNRRSSTPPQKOV 332
 QY 47 TERCSQWSPDATTLDDNGTMLPFKGEFWKSHKMDRELISER-----WKNFP---SP 96
 Db 333 PHRCTA--HFDVA--QIRGEAFFKGFYRRLTR-DRLVSLQPAQMRFRGLPHLDS 388
 QY 97 VDAAF-ROGHSYFLIKDGRVWVYPRPEKKEGYKLLQDEFPGLPPLDAVCHRGECQ 155
 Db 389 VDAVERTSDHKIVFPGDRYVWFKDNVVEGYRPRVSD-FSLPGGIDA----- 437
 QY 156 AEGVLFPGQGHNGHNGSTHGGPRVWCSPHLVLSALTSNKGATVAFSGTHWRD 215
 Db 438 ---VFSWANDR-----TYFFKQQLWRYD 459
 QY 216 --TSR--DGWHS-WPIAHQWPGPSAVDAFAFSWEE-KLYLVQGTQVYVFTKGGYTLVSG 269
 Db 460 DHTRRMDPGYPAGGPL---WRGVPSMLDDMRMSDGSATFFRG-QEYWKVLDSLELAAPG 515
 QY 270 YPRLEKE---VGTPHGILLSDVDAFICPGSSRLHIMAGRRLWLDLKSQAQATWTELP 326
 Db 516 YPGSTARMDLVGCEP-----LADAEVGPQPQ-----GR-----SGAQ----- 548
 QY 327 WPHKXVDG-ALCMKSLGPRNSCANGPGLYLHGNPNLYCYSDVEKLNAAKALPQPNVTS 385
 Db 549 ---DGLAVCS-----CTSDAHL---ALPSLLLTLP 573
 QY 386 LL 387
 Db 574 LL 575
 RESULT 6
 AAE10419
 ID AAE10419 standard; Protein; 470 AA.
 AC AAE10419;
 XX
 DT 10-DEC-2001 (first entry)
 XX
 DE Human matrix metalloproteinase-12 (MMP-12) protein.
 XX
 KM Human; matrix metalloproteinase; MMP-12; hair growth; antisense therapy;
 KM endopeptidase; skin cell; breast cancer; hair follicle; chromosome 11q22.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Peptide 1..16 Location/Qualifiers
 FT /label= Signal_peptide
 FT Protein 17..470 /label= Mature_MMP_12_protein
 FT Domain 90..96 /label= Cysteine_switch_domain
 FT Domain 161..185 /note= "Zinc and calcium binding domain"
 FT Domain

PN WO200166766-A2.
XX 13-SEP-2001.
XX 06-MAR-2001; 2001WO-US07167.
XX 06-MAR-2001; 2000US-0187196.
XX (DARW-) DARWIN MOLECULAR CORP.
XX (SCHA/) SCHATZMAN R.
XX Rajardo M, Wang K, Smith R, Moss P;
XX WPI; 2001-582276/65.
XX Novel isolated matrix metalloproteinase-25 nucleic acid molecule and
XX PT proteins encoded by them whose inhibition is useful for modulation of
XX PT hair growth in mammals -
XX
XX Example 2; Fig 3; 119pp; English.
XX
XX The present sequence is human matrix metalloproteinase (MMP)-12 protein
XX used in the exemplification of the invention. MMP-25 DNA is located
XX on chromosome 11q22. Matrix metalloproteinases are a family of zinc
XX dependent endopeptidases that function extracellularly to degrade
XX proteins typically found in the extracellular matrix. MMP-25 is expressed
XX in skin cells of mammals, particularly in breast cells and hair
XX follicles. MMP-25 DNA is useful for identifying a nucleic acid molecule
XX encoding all or part of MMP by hybridizing MMP-25 to a nucleic acid
XX sample and identifying a sequence that hybridizes in the nucleic acid
XX acid sample. The identification step involves performing polymerase
XX chain reaction (PCR) to amplify the hybridizing sequence. MMP-25
XX antibody is useful for identifying type 25 MMP. MMP-25 protein
XX inhibitors may be used to modulate hair growth and breast cancer in
XX a mammal.
XX
XX Sequence 470 AA;
SQ
Query Match 9.3%; Score 202.5; DB 22; Length 470;
Best Local Similarity 27.3%; Pred. No. 3.9e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
QY 30 SAHGNAVEGKTPDPVTE--RCSGWSFDTATLDNGNTMLFFKGEFVW---KSHKMDR 83
DB 260 SLVGDPEKRNQRLPRDPNSEPALCDPNLSFDAVTVGN-KIFFKDRFPLKVSERPTSV 318
QY 84 ELISERKKNFSPVDAFR-QGHSVFLIKGDVWVYPPPEKKEGYPKLQD-BFPGIPS 141
DB 319 NLISLWPTLPSCGIEAAVEIARNOVFLFKDDKYWLISNLRPEPNYPKSHSGFPPNVK 378
QY 142 PLDAAVECHRGECQAEGLFFQGHGHRNGTGHGNSTHGPEYKCSPHLVLSALTSDNHG 201
DB 379 KIDAAV-----FNPRFYR----- 391
QY 202 ATYAFSGTHYRLDTSRD---GHSWPIAHQWPGPSAVDAASWEK-LYLVQGTQVY 256
DB 392 -TYFFVNOYWRVYDERQMDPGYPKLITKNFGQIGPK-IDAVFYSKNKYYFFQGSNOF 449
QY 257 VF 258
DB 450 EY 451
RESULT 7
AAB84614
ID AAB84614 standard; Protein; 470 AA.
XX AAB84614;
XX
XX 05-SEP-2001 (first entry)
XX
XX Amino acid sequence of matrix metalloproteinase-12.

KW Growth factor; protein inhibitor; protease; damaged tissue;
KW platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;
KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
KW transducing growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
KW granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
KW vascular endothelial growth factor; urokinase plasminogen activator;
KW dermal ulcer; wound.
XX
XX Homo sapiens.
XX
XX WO200149309-A2.
XX 12-JUL-2001.
XX
XX 21-DEC-2000; 2000WO-IB01935.
XX
XX 29-DEC-1999; 99GB-0030768.
XX
XX (PRIZ) PRIZER LTD.
XX (PRIZ) PRIZER INC.
XX
XX Davies MJ, Huggins JP, McIntosh FS, Occlleston NL;
XX WPI; 2001-418351/44.
XX N-PSDB; AAB28229.
XX
XX Composition for the treatment of damaged tissue i.e. chronic wounds and
XX PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
XX factor -
XX
XX Disclosure; Page 560; 572pp; English.
XX
XX The specification describes a pharmaceutical composition, comprising
XX a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
XX agent inhibits the action of at least one specific adverse protein,
XX i.e. a protease, that is upregulated in a damaged tissue such as a
XX wound environment. Growth factors which are included in the composition
XX of the invention are platelet-derived growth factor (PDGF), fibroblast
XX growth factor (FGF), connective tissue derived growth factor (CTGF),
XX keratinocyte-derived growth factor (KGF), transforming growth
XX factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
XX (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
XX factor (VEGF), and chrysalin. Inhibitors which are included in the
XX composition of the invention include inhibitors of urokinase-type
XX plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
XX composition is useful for the treatment of chronic damaged tissue, i.e.
XX wounds and dermal ulcers. The present sequence represents a human MMP-12,
XX and is used to produce the composition of the invention.
XX
XX Sequence 470 AA;
SQ
Query Match 9.3%; Score 202.5; DB 22; Length 470;
Best Local Similarity 27.3%; Pred. No. 3.9e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
QY 30 SAHGNAVEGKTPDPVTE--RCSGWSFDTATLDNGNTMLFFKGEFVW---KSHKMDR 83
DB 260 SLVGDPEKRNQRLPRDPNSEPALCDPNLSFDAVTVGN-KIFFKDRFPLKVSERPTSV 318
QY 84 ELISERKKNFSPVDAFR-QGHSVFLIKGDVWVYPPPEKKEGYPKLQD-BFPGIPS 141
DB 319 NLISLWPTLPSCGIEAAVEIARNOVFLFKDDKYWLISNLRPEPNYPKSHSGFPPNVK 378
QY 142 PLDAAVECHRGECQAEGLFFQGHGHRNGTGHGNSTHGPEYKCSPHLVLSALTSDNHG 201
DB 379 KIDAAV-----FNPRFYR----- 391
QY 202 ATYAFSGTHYRLDTSRD---GHSWPIAHQWPGPSAVDAASWEK-LYLVQGTQVY 256
DB 392 -TYFFVNOYWRVYDERQMDPGYPKLITKNFGQIGPK-IDAVFYSKNKYYFFQGSNOF 449
QY 257 VF 258

Db 450 EY 451

RESULT 8

ID AAB74595 standard; Protein; 470 AA.

AC AAB74595;

DT 11-JUN-2001 (first entry)

DE Human macrophage metalloelastase HME.

KM Human; HME: macrophage metalloelastase; elastolysis; pulmonary emphysema; inflammatory disease.

OS Homo sapiens.

PN US6204043-B1.

PD 20-MAR-2001.

PF 01-MAR-1995; 95US-0396988.

PR 28-MAY-1993; 93US-0068392.

PA (UNIW) UNIV WASHINGTON.

PI Shapito SD;

DR WPI; 2001-289518/30.

DR N-PSDB; AAF81624.

PT New human macrophage metalloelastase, useful in the normal embryonic development, growth, tissue remodeling and tissue repair, particularly in studying the pathogenesis of pulmonary emphysema -

PS Claim 1; Fig 5; 20pp; English.

CC The present invention provides the protein and coding sequences of human macrophage metalloelastase (HME). The protein has elastolytic activity.

CC HME has a role in tissue remodeling and repair associated with development and inflammation, and abnormal expression can result in tumour invasiveness, arthritis and atherosclerosis. It is also thought to contribute to the pathogenesis of pulmonary emphysema and other inflammatory destructive diseases. The present sequence is the HME protein.

CC Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 22; Length 470;

Best Local Similarity 27.3%; Pred. No. 3.9e-10; Mismatches 90; Indels 63; Gaps 10;

Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGETKPDVTE--RCSDGWSPDATTLDNGTMLFFKGEFVW---KSHKMDR 83

DB 260 SLVGDPKENGRLPNPNSEPALCDPNLSFDAVTVGN-KIFFKDRFVLKVSERPKTSV 318

QY 84 ELISEERKKNPPSPVDAFR-QGHNSVFLIKGDKVWYRPEKKEGYPKLQD-EFGGIPS 141

DB 319 NLISLSMPTLPSCIEAAYBEARNQVFLFKDDKYLISNLRPEPNYPKSIHSFGFPNPFVK 378

QY 142 PLDAAYCHRGEGQAGVLFQGHGRNGTGHGNSHTHGEVYRCSPHLVLSALTSNMG 201

DB 379 KIDAAV-----FNPRFYR----- 391

QY 202 ATYAFSGTHYRLDTSRD---GHSWPIAHQWPGPSAVDAFASWEK-LYLVOGTQVY 256

DB 392 -TYFFVQNGYWRDYERQWMDPGYPLITKNFGIGPK-IDAVFYSKNKYYFFQGSNPF 449

QY 257 VF 258

Db 450 EY 451

RESULT 9

ID AAB49982 standard; Protein; 470 AA.

AC AAB49982;

DT 13-MAR-2001 (first entry)

DE Human macrophage metalloelastase.

KM Human; macrophage metalloproteinase; elastin; matrix degrading enzyme; emphysema.

OS Homo sapiens.

PN US6150152-A.

PD 21-NOV-2000.

PF 28-MAY-1993; 93US-0068392.

PR 28-MAY-1993; 93US-0068392.

PA (UNIW) UNIV WASHINGTON.

PI Shapito SD;

DR WPI; 2001-049090/06.

DR N-PSDB; AAC64994.

PT Novel human macrophage metalloelastase polynucleotides and polypeptides useful for measuring elastin degradation -

PS Claim 1; Fig 5; 20pp; English.

CC The present invention provides the coding and protein sequences for the human macrophage metalloelastase. This is a matrix degrading metalloproteinase which has the ability to degrade elastin, and can be used to measure elastin degradation.

CC Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 22; Length 470;

Best Local Similarity 27.3%; Pred. No. 3.9e-10; Mismatches 90; Indels 63; Gaps 10;

Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGETKPDVTE--RCSDGWSPDATTLDNGTMLFFKGEFVW---KSHKMDR 83

DB 260 SLVGDPKENGRLPNPNSEPALCDPNLSFDAVTVGN-KIFFKDRFVLKVSERPKTSV 318

QY 84 ELISEERKKNPPSPVDAFR-QGHNSVFLIKGDKVWYRPEKKEGYPKLQD-EFGGIPS 141

DB 319 NLISLSMPTLPSCIEAAYBEARNQVFLFKDDKYLISNLRPEPNYPKSIHSFGFPNPFVK 378

QY 142 PLDAAYCHRGEGQAGVLFQGHGRNGTGHGNSHTHGEVYRCSPHLVLSALTSNMG 201

DB 379 KIDAAV-----FNPRFYR----- 391

QY 202 ATYAFSGTHYRLDTSRD---GHSWPIAHQWPGPSAVDAFASWEK-LYLVOGTQVY 256

DB 392 -TYFFVQNGYWRDYERQWMDPGYPLITKNFGIGPK-IDAVFYSKNKYYFFQGSNPF 449

QY 257 VF 258

RESULT 10

ID AAU91061 standard; Protein; 470 AA.

AAU91061

XX AAU91061;
 AC 05-JUN-2002 (first entry)
 DT Human head and neck tumour protein, SEQ ID No 219.
 XX Human head and neck tumour protein, SEQ ID No 219.
 DE Human; head and neck cancer; tumour; cytostatic; immunogenic; vaccine.
 XX Homo sapiens.
 OS WO200212329-A2.
 XX 14-FEB-2002.
 PD 01-AUG-2001; 2001WO-US24226.
 PF 03-AUG-2000; 2000US-223281P.
 PR 16-NOV-2000; 2000US-249333P.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Fan L;
 PI WPI; 2002-257467/30.
 DR N-PSDB; ABK54023.
 XX Novel polynucleotide encoding head and neck tumour polypeptides, useful
 PT in pharmaceutical compositions, e.g. vaccines, for treating head and
 PT neck cancers -
 PS Claim 2; Page 181-182; 200pp; English.
 XX The invention relates to an isolated polynucleotide (I) comprising
 CC sequences selected from 273 sequences fully defined in the specification.
 CC (I), including its encoded polypeptide (II), an antibody binding to (II),
 CC a fusion protein comprising (II) and a T-cell population stimulated by
 CC (I) or (II) are useful for stimulating an immune response in a patient
 CC and treating head and neck cancer in a patient. An oligonucleotide (III)
 CC that hybridises to (I) is useful for determining the presence of cancer
 CC in a patient, by obtaining a biological sample from the patient,
 CC contacting the sample with (III), detecting in the sample an amount of a
 CC polynucleotide that hybridises to the oligonucleotide, and comparing the
 CC amount of polynucleotide that hybridises to the oligonucleotide to a
 CC predetermined cut-off value. (I) and (II) are useful in pharmaceutical
 CC compositions, e.g. vaccines, and other compositions for the diagnosis and
 CC treatment of head and neck cancer. AAU91061-AAU91062 represent human head
 CC and neck cancer protein sequences of the invention.
 CC
 XX Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 23; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;
 Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
 Oy 30 SAHGNVABGKTPDPVTE--RCSGWSFDTLLDNGTMLFFKGEFVW---KSHKMDR 83
 Db 260 SLYGDPRKNQRLPNDNSEPALCDPNLSFDAVTVGN-KIFFFDKRFMLKVSERPKTSV 318
 Oy 84 ELISERKKNPSPVDAAFR-QGHSVFLIKGDKVWVPEKKEKGYPKLLOD-BEPGIPS 141
 Db 319 NLISLMPFLPSGIEAAVEIARNGVFLFKDKYWLISNLRPEPNYPKSHSPGPNFVK 378
 Oy 142 PLDAAVECHRGCECAEGVLPFOGHGHRNGTGHGNSTHHGEYMRCSPHLVLSALTSNMG 201
 Db 379 KIDAAV-----FNPRFYR----- 391
 Oy 202 ATYAFSGTHWRDLTSD---GHSWPIAHQWPGSAPVDAAFSMEK-LYLVGQTGVY 256
 Db 392 -TYFFVNDQYWRBYDERQMDPBGYPKLTITKXFGIGPK-IDAVFSKXKYYTFFQGSNOF 449
 Oy 257 VF 258

Db 450 EY 451
 RESULT 11
 ID ABP96800 standard; Protein, 470 AA.
 XX ABP96800;
 AC 05-JUN-2003 (first entry)
 DT Human COPD related protein SEQ ID NO:50.
 XX Human COPD related protein SEQ ID NO:50.
 DE Human; chronic obstructive pulmonary disease; COPD; chronic lung disease.
 XX Homo sapiens.
 OS WO200297127-A2.
 XX 05-DEC-2002.
 PD 28-MAY-2002; 2002WO-EP05835.
 PF 31-MAY-2001; 2001GB-0013266.
 PR (FARB) BAYER AG.
 PA Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;
 PI WPI; 2003-140492/13.
 DR N-PSDB; ACC46771.
 XX Predicting, diagnosing or prognosing chronic lung disease, by detecting
 PT a chronic obstructive pulmonary disease (COPD) gene in a biological
 PT sample -
 PS Claim 8; Page 153-155; 214pp; English.
 XX The present invention describes a method for predicting, diagnosing or
 CC prognosing chronic lung disease by detecting a chronic obstructive
 CC pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
 CC ACC46777, which encode the COPD related proteins in ABP96779 to
 CC ABP96806). The method is useful for predicting, diagnosing or prognosing
 CC chronic lung disease in a biological sample. The COPD genes and proteins
 CC encoded by them from the present invention (I) can be used for treating
 CC or preventing chronic lung disease in a mammal. (I) can be used in an
 CC animal model for determining the efficacy, toxicity, or side effects of
 CC treatment with (I), and determining the mechanism of action of (I).
 CC ACC46778 to ACC46903 represent COPD related PCR primers and probes used
 CC in an example from the present invention.
 CC
 XX Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 24; Length 470;
 Best Local Similarity 27.3%; Pred. No. 3.9e-10;
 Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
 Oy 30 SAHGNVABGKTPDPVTE--RCSGWSFDTLLDNGTMLFFKGEFVW---KSHKMDR 83
 Db 260 SLYGDPRKNQRLPNDNSEPALCDPNLSFDAVTVGN-KIFFFDKRFMLKVSERPKTSV 318
 Oy 84 ELISERKKNPSPVDAAFR-QGHSVFLIKGDKVWVPEKKEKGYPKLLOD-BEPGIPS 141
 Db 319 NLISLMPFLPSGIEAAVEIARNGVFLFKDKYWLISNLRPEPNYPKSHSPGPNFVK 378
 Oy 142 PLDAAVECHRGCECAEGVLPFOGHGHRNGTGHGNSTHHGEYMRCSPHLVLSALTSNMG 201
 Db 379 KIDAAV-----FNPRFYR----- 391
 Oy 202 ATYAFSGTHWRDLTSD---GHSWPIAHQWPGSAPVDAAFSMEK-LYLVGQTGVY 256
 Db 392 -TYFFVNDQYWRBYDERQMDPBGYPKLTITKXFGIGPK-IDAVFSKXKYYTFFQGSNOF 449

QY 257 VF 258
 DB 450 EY 451

RESULT 12
 ABR48203
 ID ABR48203 standard; Protein; 470 AA.
 XX
 AC ABR48203;
 XX
 DT 12-JUN-2003 (first entry)
 XX
 DE Human bladder cancer associated protein sequence SEQ ID NO:125.
 XX
 KM Human, bladder cancer; cytostatic; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003003906-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 03-JUL-2002; 2002WO-US21338.
 XX
 PR 03-JUL-2001; 2001US-302814P.
 XX
 PR 03-AUG-2001; 2001US-310099P.
 PR 08-NOV-2001; 2001US-343705P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 12-APR-2002; 2002US-372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Mack DH, Aziz N;
 XX
 DR WPI; 2003-201532/19.
 DR N-PSDB; ACC51017.
 XX
 PT Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprising contacting a biological sample from the patient with
 PT a bladder cancer-associated polynucleotide or antibody -
 XX
 PS Claim 10; Page 275; 307pp; English.
 XX
 CC The present invention describes a method for detecting a bladder cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridizes to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications.
 CC
 SQ Sequence 470 AA;
 XX
 QY Query Match 9.3%; Score 202.5; DB 24; Length 470;
 DB Best Local Similarity 27.3%; Pred. No. 3.9e-10;
 DB Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
 QY 30 SAHGNVAGETKDDPDVTE--RSGDGSFPATYLLDNGTLMFKGGEFVW-----KSHKMDR 83
 DB 260 SLVGDPKRNQRLPNPNDSEBALCPNLSFPAVTVGN-KLFFKDRFWMKVSBRPTSV 318
 QY 84 ELISERKKNPPSPVDAFR-QGHNVSFLIGDKWKVTPPEKKEGKYLQD-EFPGIIPS 141
 DB 319 NLISLWPTLPSGIEAAYEIEARQVPLFKDDKXWMLISNLRPEPNYKSIHSFGPFNVK 378

QY 142 PLDAVECHRGECQAGVLFFQGHGRNCTGHGNSTHGPEYKRCSPHLVLSALTSDNHG 201
 DB 379 KIDAAV-----FNDRFYR----- 391

QY 202 ATYAFSGTHYWRUDTSRD-----GHSWPIAHQWPGSPSAVDAAFSWEK-LYLVQSTQVY 256
 DB 392 -TYFVDNQWYRDERRQWMDPGYPLITKNFGIGPK-IDAVFYSKNRYFFQGSNPF 449

QY 257 VF 258
 DB 450 EY 451

RESULT 13
 ABUS6421
 ID ABUS6421 standard; Protein; 470 AA.
 XX
 AC ABUS6421;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #14.
 XX
 KM Lung cancer-associated polypeptide; cytostatic; emphysema;
 KM antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KM chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KM interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 PN WO200286443-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 18-APR-2002; 2002WO-US12476.
 XX
 PR 18-APR-2001; 2001US-284770P.
 PR 10-MAY-2001; 2001US-290492P.
 PR 09-NOV-2001; 2001US-339245P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-354370P.
 PR 12-APR-2002; 2002US-372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N, Murray R;
 XX
 DR WPI; 2003-093161/08.
 DR N-PSDB; ABX76137.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer -
 XX
 PS Claim 27; Page 198; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial

CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
CC polypeptides of the invention.

XX Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 24; Length 470;
Best Local Similarity 27.3%; Pred. No. 3.9e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGTKPPDVT--RCSGWSFDATTLDDNGTMLFKGEFVW---KSHKMDR 83
DB 260 SLVGDPEKNQRLPNDSEPALCDPNLSFDAVTVGN-KIFFKDRFWMKVSRPKTSV 318
QY 84 ELISERWKNPPSPVDAFR-QGHSNVLKGDKVVVYPEKKEKGYPLLOD-EFGIGPS 141
DB 319 NLISLWPTLPSGIEAAVEIARQVFLPKDKYWLISNLRPEPNYPSKISFGFPNFKV 378
QY 142 PLDAVECHRGCEQAEGLFFQGHGHRNGTGHGNSHTHGPRVMCSPHLVLSALTSNMG 201
DB 379 KIDAAV-----FNPRPYR----- 391
QY 202 ATYAFSGTHWRLDTSRD---GWSHWPDAHQMFGPSAVDAAFSWEK-LYLVQGTQVY 256
DB 392 -TYFVDNQWRYRDERRQMDPGYPLKLTKNFGIGRK-IDAVFYSKXKYTYFFQGSNQF 449
QY 257 VF 258
DB 450 EY 451

RESULT 14

ID ABUS6663 standard; Protein; 470 AA.

AC ABUS6663;

DT 02-APR-2003 (first entry)

XX Lung cancer-associated polypeptide #256.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;
XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Undenified.

OS WO200286443-A2.

PN 31-OCT-2002.

PD 18-APR-2002; 2002WO-US12476.

PE 18-APR-2001; 2001US-284770P.

PR 10-MAY-2001; 2001US-290492P.

PR 09-NOV-2001; 2001US-339245P.

PR 13-NOV-2001; 2001US-350666P.

PR 29-NOV-2001; 2001US-334370P.

PR 12-APR-2002; 2002US-372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX MPI: 2003-093161/08.

XX N-PBDB; ABX76392.

PT Detecting a lung cancer-associated transcript in a cell from a patient
for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer
XX
XX
PS Claim 27; Page 385-386; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
XX transcript in a cell from a patient, comprising contacting a biological
XX sample from the patient with a polynucleotide that selectively hybridizes
XX to a sequence that is at least 80 % identical to a gene that exhibits
XX increased or decreased expression in lung cancer samples. Lung
XX cancer-associated polynucleotides and polypeptides are used for
XX identifying a compound that modulates a lung cancer-associated
XX polypeptide, for inhibiting proliferation of a lung cancer-associated
XX cell to treat lung cancer in a patient and for treating a mammal having
XX lung cancer by administering a modulatory compound identified. The
XX methods are useful for treating lung cancer, such as small cell lung
XX cancer, non-small cell lung cancer or other benign or precancerous
XX lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
XX pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
XX pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
XX and polypeptides are useful for diagnostic purposes and as targets for
XX screening for therapeutic compounds that modulate lung cancer, such as
XX antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
XX polypeptides of the invention.

XX Sequence 470 AA;

Query Match 9.3%; Score 202.5; DB 24; Length 470;
Best Local Similarity 27.3%; Pred. No. 3.9e-10;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGTKPPDVT--RCSGWSFDATTLDDNGTMLFKGEFVW---KSHKMDR 83
DB 260 SLVGDPEKNQRLPNDSEPALCDPNLSFDAVTVGN-KIFFKDRFWMKVSRPKTSV 318
QY 84 ELISERWKNPPSPVDAFR-QGHSNVLKGDKVVVYPEKKEKGYPLLOD-EFGIGPS 141
DB 319 NLISLWPTLPSGIEAAVEIARQVFLPKDKYWLISNLRPEPNYPSKISFGFPNFKV 378
QY 142 PLDAVECHRGCEQAEGLFFQGHGHRNGTGHGNSHTHGPRVMCSPHLVLSALTSNMG 201
DB 379 KIDAAV-----FNPRPYR----- 391
QY 202 ATYAFSGTHWRLDTSRD---GWSHWPDAHQMFGPSAVDAAFSWEK-LYLVQGTQVY 256
DB 392 -TYFVDNQWRYRDERRQMDPGYPLKLTKNFGIGRK-IDAVFYSKXKYTYFFQGSNQF 449
QY 257 VF 258
DB 450 EY 451

RESULT 15

ID AAB43772 standard; Protein; 473 AA.

AC AAB43772;

DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1217.

XX Human, cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiviral;
XX antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
XX dermatologic; neuoprotective; thrombolytic; coagulant; nootropic;
XX vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 16, 2003, 06:46:44 ; Search time 28 Seconds
(without alignments)
590.841 Million cell updates/sec

Title: US-09-900-448-2
Perfect score: 2185
Sequence: 1 MARVLGAPVALGLMSLWSL.....NAKALPQPVNTSLGCTH 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	202.5	9.3	470	US-08-068-392-2	Sequence 2, Appl1
2	202.5	9.3	470	US-08-396-988-2	Sequence 2, Appl1
3	202.5	9.3	470	US-09-391-104-26	Sequence 26, Appl1
4	199	9.1	517	US-09-391-104-32	Sequence 32, Appl1
5	199	9.1	519	US-09-211-704A-7	Sequence 7, Appl1
6	189.5	8.7	476	US-08-704-711A-21	Sequence 21, Appl1
7	189.5	8.7	476	US-08-448-489-14	Sequence 14, Appl1
8	189.5	8.7	476	US-09-521-220-21	Sequence 21, Appl1
9	189.5	8.7	476	US-09-391-104-22	Sequence 22, Appl1
10	183.5	8.4	444	US-09-178-002-2	Sequence 2, Appl1
11	183.5	8.4	467	US-09-391-104-24	Sequence 4, Appl1
12	183.5	8.4	467	US-08-448-489-13	Sequence 13, Appl1
13	183.5	8.4	468	US-08-448-489-13	Sequence 23, Appl1
14	181.5	8.3	492	US-07-794-393-4	Sequence 4, Appl1
15	181.5	8.3	492	US-08-001-711-4	Sequence 4, Appl1
16	181.5	8.3	488	US-07-794-393-2	Sequence 2, Appl1
17	179.5	8.2	488	US-08-001-711-2	Sequence 2, Appl1
18	179.5	8.2	488	US-08-704-711A-22	Sequence 22, Appl1
19	179.5	8.2	488	US-09-521-220-22	Sequence 22, Appl1
20	179.5	8.2	488	US-09-391-104-31	Sequence 31, Appl1
21	179.5	8.2	489	US-08-448-489-11	Sequence 11, Appl1
22	179.5	8.2	489	US-08-448-489-11	Sequence 20, Appl1
23	177	8.1	477	US-08-704-711A-20	Sequence 15, Appl1
24	177	8.1	477	US-08-448-489-15	Sequence 9, Appl1
25	177	8.1	477	US-08-281-313-1	Sequence 20, Appl1
26	177	8.1	477	US-09-521-220-20	Sequence 21, Appl1
27	177	8.1	477	US-09-391-104-21	

28	176.5	8.1	469	3	US-08-704-711A-16	Sequence 16, Appl1
29	176.5	8.1	469	3	US-08-448-489-12	Sequence 12, Appl1
30	176.5	8.1	469	4	US-09-521-220-16	Sequence 16, Appl1
31	176	8.1	471	4	US-09-391-104-25	Sequence 25, Appl1
32	175	8.0	466	3	US-08-704-711A-17	Sequence 17, Appl1
33	175	8.0	466	4	US-09-521-220-17	Sequence 17, Appl1
34	172	7.9	664	3	US-09-211-704A-8	Sequence 8, Appl1
35	172	7.9	669	3	US-08-704-711A-8	Sequence 3, Appl1
36	172	7.9	669	4	US-09-521-220-3	Sequence 3, Appl1
37	172	7.9	669	4	US-09-391-104-29	Sequence 29, Appl1
38	169.5	7.8	508	4	US-09-391-104-18	Sequence 18, Appl1
39	167.5	7.7	411	4	US-09-171-545-3	Sequence 3, Appl1
40	167.5	7.7	416	4	US-09-171-545-4	Sequence 4, Appl1
41	167.5	7.7	508	4	US-09-171-545-1	Sequence 1, Appl1
42	164.5	7.5	324	2	US-08-816-755-2	Sequence 2, Appl1
43	164.5	7.5	324	3	US-09-090-673-2	Sequence 2, Appl1
44	161.5	7.4	532	3	US-09-294-841-2	Sequence 2, Appl1
45	154	7.0	419	4	US-09-171-545-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-08-068-392-2
; Sequence 2, Application US/08068392
; Patent No. 6150152
; GENERAL INFORMATION:
; APPLICANT: Shaprio, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A35M
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,392
; FILING DATE: 19930528
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25275
; REFERENCE/DOCKET NUMBER: 07-24(12406)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-068-392-2

Query Match 9.3%; Score 202.5; DB 3; Length 470;
Best Local Similarity 27.3%; Pred. No. 2.4e-12;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;

QY 30 SAHGNVAEGETKPPDYTE--RCSDGSPDATTDDNGTMLFFKGEFVW----KSHKWR 83
DB 260 SLYDPPKRNQGLRPDPNSRPLCDPNLSFDVAVTVGN-KIFFPDRFMLKVSRRPTSV 318
QY 84 ELISERWKNFSPVDAAFR-QGHSVFLIKDKKRWVVPPEKKEGYPLLOD-EFPGIPS 141
DB 319 NLISLWPTLPSGIEAAVIEARNOVFLFKDKKRWLISNLRPBNYPKSISFGFPMFVK 378

QY 142 PLDAVECHRGCOAGVLFPOGHGHRNGTGHGNSTHGPEYMRCSPHVLVLSALTSDNHG 201
| : : : : :
Db 379 KIDAAV-----FNRRFYR----- 391
QY 202 ATYAFSGTHYWRDLTSDH-----GHSWPIAHQWPOGSAVDAAFSWEK-LYLVQGTQVY 256
| : : : : :
Db 392 -TYFFVNDQYWRDYDERQWMDPGYPKLITKNQIGPK-IDAVFYSKNKYFFFGQSNQF 449
QY 257 VF 258
| : : : : :
Db 450 EY 451
RESULT 2
US-08-396-988-2
; Sequence 2, Application US/08396988
; Patent No. 6204043
; GENERAL INFORMATION:
; APPLICANT: Shapito, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,988
; FILING DATE: 01-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,392
; FILING DATE: 28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25275
; REFERENCE/DOCKET NUMBER: 07-24(12406)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-396-988-2
Query Match 9.3%; Score 202.5; DB 3; Length 470;
Best Local Similarity 27.3%; Pred. No. 2.4e-12;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
QY 30 SAHGNAVEGETKPDPPVTE--RCSDGWSPDATLTDNGTMLFPKGEFVW-----KSHKMDR 83
| : : : : :
Db 260 SLVGDPKRNQRLNPNNSPALCDPNLSFDAVTTVGN-KIFFKDRFPLKXSERKTSV 318
QY 84 ELISERKNFPPSPVDAFR-QGHSNVFLIKGDKWVYPPKKEKGYPKLQD-EFPGIPS 141
| : : : : :
Db 319 NLISSLMPTLPISGIEAAYEIEARNQVFLFKDXYWLSNLRBPENYFKSIHSFGFPNFK 378
QY 142 PLDAVECHRGCOAGVLFPOGHGHRNGTGHGNSTHGPEYMRCSPHVLVLSALTSDNHG 201
| : : : : :
Db 379 KIDAAV-----FNRRFYR----- 391
QY 202 ATYAFSGTHYWRDLTSDH-----GHSWPIAHQWPOGSAVDAAFSWEK-LYLVQGTQVY 256
| : : : : :
Db 450 EY 451

Db 392 -TYFFVNDQYWRDYDERQWMDPGYPKLITKNQIGPK-IDAVFYSKNKYFFFGQSNQF 449
QY 257 VF 258
| : : : : :
Db 450 EY 451
RESULT 3
US-09-391-104-26
; Sequence 26, Application US/09391104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-391-104-26
Query Match 9.3%; Score 202.5; DB 4; Length 470;
Best Local Similarity 27.3%; Pred. No. 2.4e-12;
Matches 66; Conservative 23; Mismatches 90; Indels 63; Gaps 10;
QY 30 SAHGNAVEGETKPDPPVTE--RCSDGWSPDATLTDNGTMLFPKGEFVW-----KSHKMDR 83
| : : : : :
Db 260 SLVGDPKRNQRLNPNNSPALCDPNLSFDAVTTVGN-KIFFKDRFPLKXSERKTSV 318
QY 84 ELISERKNFPPSPVDAFR-QGHSNVFLIKGDKWVYPPKKEKGYPKLQD-EFPGIPS 141
| : : : : :
Db 319 NLISSLMPTLPISGIEAAYEIEARNQVFLFKDXYWLSNLRBPENYFKSIHSFGFPNFK 378
QY 142 PLDAVECHRGCOAGVLFPOGHGHRNGTGHGNSTHGPEYMRCSPHVLVLSALTSDNHG 201
| : : : : :
Db 379 KIDAAV-----FNRRFYR----- 391
QY 202 ATYAFSGTHYWRDLTSDH-----GHSWPIAHQWPOGSAVDAAFSWEK-LYLVQGTQVY 256
| : : : : :
Db 392 -TYFFVNDQYWRDYDERQWMDPGYPKLITKNQIGPK-IDAVFYSKNKYFFFGQSNQF 449
QY 257 VF 258
| : : : : :
Db 450 EY 451
RESULT 4
US-09-391-104-32
; Sequence 32, Application US/09391104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07

PRIOR APPLICATION NUMBER: US 08/814,394
 PRIOR FILING DATE: 1997-03-11
 NUMBER OF SEQ ID NOS: 35
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO: 32
 LENGTH: 517
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-391-104-32

Query March 9.1%; Score 199; DB 4; Length 517;
 Best Local Similarity 25.0%; Pred. No. 6,3e-12;
 Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

5 LGAPVALGL-----WSLCMSLAATPL-----PTSAHGNVAEGETKPDPTVTERCS 51
 188 VGDPLRYGLPEYDKRVWQLYGVRESVPTAQPEEPPLPEPPDNRSAPPRKDVPHRCS 247
 52 DGMSFDATLDDNGTMTLFFKGEFVWKSXKMDRELISER-----WKNFP---SPVDAAF 101
 248 T--HFDAVA-QIRGEAFPFKGYFWRLTR-DRHLVSLQPAQMRFRWGLPLHLDSVDVAVY 303
 102 -RQGHNSVFLIKGDKVWVYPPPEKKEKGYPKLLQDEFPGIPSPPLDAAVECHRGECQAGSVL 160
 304 ERTSDHKIVFFKGRYVFKDNVVEEGYPRPVSD-FSLPPGCGIDAA----- 348
 161 FFGGHRNGTGHGNSTHGPEVWRCSPHLVLSALTSNMGATYAFSGTHYWRDLT--SRD 219
 349 FSWAHNDR-----TYFFKDQLYWRVDHTRH 374
 220 GMSWPI-AHQWPGPSAVDAAFSWEK-KLYLVQGTQVYVFLTKGYTLVSGYPRLEKE 277
 375 MDGYPAGSPLMRGVPSLDDAMRWSDGASYFFRG-QEYWKVLDGELEVAAGYQSTARD 433
 278 -----VGTPHGIILDSVDAA 292
 434 WLVCQDSQADGSVAAGVDAA 453

RESULT 5

US-09-211-704A-7
 Sequence 7, Application US/09211704A
 Patent No. 6271014
 GENERAL INFORMATION:
 APPLICANT: de Saint-Vis, Blandine Marie
 APPLICANT: Fossiez, Francois
 APPLICANT: Caux, Christophe
 APPLICANT: Lebecque, Serge J.E.
 TITLE OF INVENTION: Mammalian Proteinases; Related Reagents
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DNAX Research Institute
 STREET: 901 California Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/211.704A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/005,263
 FILING DATE: 09-JAN-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Ching, Edwin P.
 REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: SF0781K
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 852-9196
 TELEFAX: (650) 496-1200
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 519 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-211-704A-7

Query March 9.1%; Score 199; DB 3; Length 519;
 Best Local Similarity 25.0%; Pred. No. 6,4e-12;
 Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

5 LGAPVALGL-----WSLCMSLAATPL-----PTSAHGNVAEGETKPDPTVTERCS 51
 190 VGDPLRYGLPEYDKRVWQLYGVRESVPTAQPEEPPLPEPPDNRSAPPRKDVPHRCS 249
 52 DGMSFDATLDDNGTMTLFFKGEFVWKSXKMDRELISER-----WKNFP---SPVDAAF 101
 250 T--HFDAVA-QIRGEAFPFKGYFWRLTR-DRHLVSLQPAQMRFRWGLPLHLDSVDVAVY 305
 102 -RQGHNSVFLIKGDKVWVYPPPEKKEKGYPKLLQDEFPGIPSPPLDAAVECHRGECQAGSVL 160
 306 ERTSDHKIVFFKGRYVFKDNVVEEGYPRPVSD-FSLPPGCGIDAA----- 350
 161 FFGGHRNGTGHGNSTHGPEVWRCSPHLVLSALTSNMGATYAFSGTHYWRDLT--SRD 219
 351 FSWAHNDR-----TYFFKDQLYWRVDHTRH 376
 220 GMSWPI-AHQWPGPSAVDAAFSWEK-KLYLVQGTQVYVFLTKGYTLVSGYPRLEKE 277
 377 MDGYPAGSPLMRGVPSLDDAMRWSDGASYFFRG-QEYWKVLDGELEVAAGYQSTARD 435
 278 -----VGTPHGIILDSVDAA 292
 436 WLVCQDSQADGSVAAGVDAA 455

RESULT 6

US-08-704-711A-21
 Sequence 21, Application US/08704711A
 Patent No. 6114159
 GENERAL INFORMATION:
 APPLICANT: WILZ, Horst
 APPLICANT: HINZMANN, Bernd
 TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/704,711A
 FILING DATE: 20-NOV-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/DE95/00357
 FILING DATE: 17-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 4438838.1

FILING DATE: 21-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4409663.1
FILING DATE: 17-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 26083/124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-704-711A-21

Query Match 8.7%; Score 189.5; DB 3; Length 476;
Best Local Similarity 28.4%; Pred. No. 5.4e-11;

Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

Qy 25 PLPPTSAHGVNAGETKPDVTERCSDGWSFDA-TTLDNGTMTLFFKGFVWKSHKMDR 83
Db 272 PLVPTK---SVPGSEMP-----AKCDPALSPDAISTL--RGEYLFKDRYFWRSHMNP 321

Qy 84 E-----LISERWKVFPSPVDAAFR-QGHNVSFLIKGDKVWV-----YP 120
Db 322 EPEFHLSAFWPSLPSTLDAAYEVNSRDVTFIKGNEFWAIRGNEVOAGYPRGIHTLGF 381

Qy 121 P-----EKK-----EKGYPKLQDEFGIPSPDLAAVE 148
Db 382 PTIRKIDAAVSDKEKKTYFFAADKTYRFDENSQSMEOGFPRLIADFPGEVPEKVDVAVL- 440

Qy 149 CHRGECCAGVLFF 162
Db 441 -----QAFGFFYF 448

RESULT 7
US-08-448-489-14
Sequence 14, Application US/08448489
Patent No. 6184022
GENERAL INFORMATION:
APPLICANT: SEIKI, Motoharu
APPLICANT: SATO, Hiroshi
APPLICANT: SHINAGAWA, Akira
TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
FILE REFERENCE: 55-280P
CURRENT APPLICATION NUMBER: US/08/448,489
CURRENT FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 14
LENGTH: 476
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Known Member of
OTHER INFORMATION: Matrix Metalloproteinase Family
US-08-448-489-14

Query Match 8.7%; Score 189.5; DB 3; Length 476;
Best Local Similarity 28.4%; Pred. No. 5.4e-11;

Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

Qy 25 PLPPTSAHGVNAGETKPDVTERCSDGWSFDA-TTLDNGTMTLFFKGFVWKSHKMDR 83
Db 272 PLVPTK---SVPGSEMP-----AKCDPALSPDAISTL--RGEYLFKDRYFWRSHMNP 321

Qy 84 E-----LISERWKVFPSPVDAAFR-QGHNVSFLIKGDKVWV-----YP 120

Db 322 EPEFHLSAFWPSLPSTLDAAYEVNSRDVTFIKGNEFWAIRGNEVOAGYPRGIHTLGF 381
Qy 121 P-----EKK-----EKGYPKLQDEFGIPSPDLAAVE 148
Db 382 PTIRKIDAAVSDKEKKTYFFAADKTYRFDENSQSMEOGFPRLIADFPGEVPEKVDVAVL- 440
Qy 149 CHRGECCAGVLFF 162
Db 441 -----QAFGFFYF 448

RESULT 8
US-09-521-220-21
Sequence 21, Application US/09521220
Patent No. 639348
GENERAL INFORMATION:
APPLICANT: WILU, Horst
HINZMANN, Bernd
TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
METALLOPROTEASES, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/521,220
FILING DATE: 08-Mar-2000
CLASSIFICATION: <Unknown>
21-OCT-1994
17-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/704,711
FILING DATE: <Unknown>
APPLICATION NUMBER: DE 4438838.1
FILING DATE: 21-OCT-1994
APPLICATION NUMBER: DE 4409663.1
FILING DATE: 17-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 26083/124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-521-220-21

Query Match 8.7%; Score 189.5; DB 4; Length 476;
Best Local Similarity 28.4%; Pred. No. 5.4e-11;

Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

Qy 25 PLPPTSAHGVNAGETKPDVTERCSDGWSFDA-TTLDNGTMTLFFKGFVWKSHKMDR 83
Db 272 PLVPTK---SVPGSEMP-----AKCDPALSPDAISTL--RGEYLFKDRYFWRSHMNP 321

Qy 84 E-----LISERWKVFPSPVDAAFR-QGHNVSFLIKGDKVWV-----YP 120

Db 322 EPEFLISAFWPSLPSYLDAAVEVNSRDTVFIFKGNEMFAIRGNEVQAGYPRGHTLGF 381
Qy 121 P-----EKK-----EKGYPLDDEFGPISPLDAVE 148
Db 382 PTKRIDAAVSDKEKKTYFFPAADKYRWFENDSOGFPLLIADDFGVPEPKVDALV 440
Qy 149 CHRGCOAEGVLFF 162
Db 441 -----QAFGFYF 448

RESULT 9
US-09-391-104-22

; Sequence 22, Application US/09391104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073. US. P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-391-104-22

Query Match 8.7%; Score 189.5; DB 4; Length 476;
Best Local Similarity 28.4%; Pred. No. 5.4e-11;
Matches 55; Conservative 24; Mismatches 42; Indels 73; Gaps 10;

Qy 25 PLPPTSAHGNVAEGTKPDPTVTERCSDGWSFDA-TTLDNGTMLFKGEEFVWKS HKWDR 83
Db 272 PLVPTK---SVSGSEMP-----AKCDPALSFDAISTL--RGEYLFKDRYFWRSHNRP 321
Qy 84 E---LISERKNPSPVDAAR-OGHNSVFLIKDKYV-----YP 120
Db 322 EPEFLISAFWPSLPSYLDAAVEVNSRDTVFIFKGNEMFAIRGNEVQAGYPRGHTLGF 381
Qy 121 P-----EKK-----EKGYPLDDEFGPISPLDAVE 148
Db 382 PTKRIDAAVSDKEKKTYFFPAADKYRWFENDSOGFPLLIADDFGVPEPKVDALV 440
Qy 149 CHRGCOAEGVLFF 162
Db 441 -----QAFGFYF 448

RESULT 10
US-09-178-002-2

; Sequence 2, Application US/09178002
; Patent No. H001973
; GENERAL INFORMATION:
; APPLICANT: Hu, Shou-Ih
; TITLE OF INVENTION: Human Neutrophil Collagenase Splice Variant
; FILE REFERENCE: CGC 2048
; CURRENT APPLICATION NUMBER: US/09/178,002
; CURRENT FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 444
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-178-002-2

Query Match 8.4%; Score 183.5; DB 1; Length 444;
Best Local Similarity 24.4%; Pred. No. 2e-10;
Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;

Qy 23 ATPPPTSAHGNVAEGTKPDPTVTERCSDGWSFDA-TTLDNGTMLFKGEEFVWKS HKWDR 79
Db 242 SNPIQPT-----GPSTPKP-----CDPSLTFDAITTL--RGEILFFKDRYFWRSHNRP 286
Qy 80 --KMDRELISERKNPSPVDAARQ-CHNSVFLIKDKYVVPPEKKEKGYPLDOD-E 135
Db 287 LQRYENMFISLFWPSLPTGIGQAYEDFDRLIFLFGKNGYWALSGLYDILQGYPRDISNYG 346
Qy 136 FPGIPSPDLAAVECHRGCOAEGVLFPQGHGHRNGTGCHNSTHGPYMRCSPLVLAL 195
Db 347 FPGSVQALDAV-----FYR----- 361
Qy 196 TSDNHGATVAFSGTHYWRDLTSR---DGMHSWPIAHQWPGPSAVDAAFSMEKLYLVQG 252
Db 362 ----SKTYFVNDQFWRYNQRFQFMEPGYKPSISGAFPGIESKVDVAFQOEHFHFVPSG 416
Qy 253 TOYTVF 258
Db 417 PRYVAF 422

RESULT 11
US-09-178-002-4

; Sequence 4, Application US/09178002
; Patent No. H001973
; GENERAL INFORMATION:
; APPLICANT: Hu, Shou-Ih
; TITLE OF INVENTION: Human Neutrophil Collagenase Splice Variant
; FILE REFERENCE: CGC 2048
; CURRENT APPLICATION NUMBER: US/09/178,002
; CURRENT FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-178-002-4

Query Match 8.4%; Score 183.5; DB 1; Length 467;
Best Local Similarity 24.4%; Pred. No. 2.2e-10;
Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;

Qy 23 ATPPPTSAHGNVAEGTKPDPTVTERCSDGWSFDA-TTLDNGTMLFKGEEFVWKS HKWDR 79
Db 265 SNPIQPT-----GPSTPKP-----CDPSLTFDAITTL--RGEILFFKDRYFWRSHNRP 309
Qy 80 --KMDRELISERKNPSPVDAARQ-CHNSVFLIKDKYVVPPEKKEKGYPLDOD-E 135
Db 310 LQRYENMFISLFWPSLPTGIGQAYEDFDRLIFLFGKNGYWALSGLYDILQGYPRDISNYG 369
Qy 136 FPGIPSPDLAAVECHRGCOAEGVLFPQGHGHRNGTGCHNSTHGPYMRCSPLVLAL 195
Db 370 FPGSVQALDAV-----FYR----- 384
Qy 196 TSDNHGATVAFSGTHYWRDLTSR---DGMHSWPIAHQWPGPSAVDAAFSMEKLYLVQG 252
Db 385 ----SKTYFVNDQFWRYNQRFQFMEPGYKPSISGAFPGIESKVDVAFQOEHFHFVPSG 439
Qy 253 TOYTVF 258
Db 440 PRYVAF 445

RESULT 12
US-09-391-104-24

```
Sequence 24, Application US/09391104
Patent No. 6399371
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Falduto, Michael T.
APPLICANT: Magnuson, Scott R.
APPLICANT: Morgan, Douglas W.
TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
TITLE OF INVENTION: OF USING SAME
FILE REFERENCE: 6073.US.P1
CURRENT APPLICATION NUMBER: US/09/391,104
CURRENT FILING DATE: 1999-09-07
PRIOR APPLICATION NUMBER: US 08/814,394
PRIOR FILING DATE: 1997-03-11
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens
US-09-391-104-24
```

```
Query Match 8.4%; Score 183.5; DB 4; Length 467;
Best Local Similarity 24.4%; Pred. No. 2.2e-10;
Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;
```

```
QY 23 ATPLPPTSAHGNVAEETKDPDPTERCSDGMSFDA-TTLDNGTMLFFKGEFVWKS-- 79
DB 265 SNPIQPT-----GPTSKP-----CDPSLTFTDAITTL--RGEILFFDRYFWRHPQ 309
QY 80 --KMDRELISERKWNKPSPYDAAFRO-GHNSVFLIKGDKVWVYPRPEKKEGYPLQD-E 135
DB 310 LQREVMNFIISLFWPSLPTGTGQAAYEDPDRDLIFLFGKNQWALSGYDILQGYPKDISNG 369
QY 136 FPGIPSPDLAAVECHRGCEQAEGLVFPQGHGRNGTGHNSTHGPEYRCSPHLVLSAL 195
DB 370 FPGSVQALDAV-----FYR----- 384
QY 196 TSDNHGATTAFSGTHVWRDLTSR---DGMHSMPIAHQWPGSPAVDAAFSWEKLYLVQG 252
DB 385 -----SKTYFFVNDQFWRVNDQRFMEPGYPKISGAFPGIESKVDVAFQOEHPFHVFG 439
QY 253 TOYVVF 258
DB 440 PRTYAF 445
```

RESULT 13

```
US-08-448-489-13
Sequence 13, Application US/08448489
Patent No. 6184022
GENERAL INFORMATION:
APPLICANT: SEIKI, Motoharu
APPLICANT: SATO, Hiroshi
APPLICANT: SHINAGAWA, Akira
TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
FILE REFERENCE: 55-290P
CURRENT APPLICATION NUMBER: US/08/448,489
CURRENT FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 468
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: X = UNKNOWN
OTHER INFORMATION: Description of Unknown Organism: Known Member of
OTHER INFORMATION: Matrix Metalloproteinase Family
US-08-448-489-13
```

```
Query Match 8.4%; Score 183.5; DB 3; Length 468;
Best Local Similarity 24.4%; Pred. No. 2.2e-10;
Matches 60; Conservative 31; Mismatches 80; Indels 75; Gaps 10;
```

```
QY 23 ATPLPPTSAHGNVAEETKDPDPTERCSDGMSFDA-TTLDNGTMLFFKGEFVWKS-- 79
DB 265 SNPIQPT-----GPTSKP-----CDPSLTFTDAITTL--RGEILFFDRYFWRHPQ 309
QY 80 --KMDRELISERKWNKPSPYDAAFRO-GHNSVFLIKGDKVWVYPRPEKKEGYPLQD-E 135
DB 310 LQREVMNFIISLFWPSLPTGTGQAAYEDPDRDLIFLFGKNQWALSGYDILQGYPKDISNG 369
QY 136 FPGIPSPDLAAVECHRGCEQAEGLVFPQGHGRNGTGHNSTHGPEYRCSPHLVLSAL 195
DB 370 FPGSVQALDAV-----FYR----- 384
QY 196 TSDNHGATTAFSGTHVWRDLTSR---DGMHSMPIAHQWPGSPAVDAAFSWEKLYLVQG 252
DB 385 -----SKTYFFVNDQFWRVNDQRFMEPGYPKISGAFPGIESKVDVAFQOEHPFHVFG 439
QY 253 TOYVVF 258
DB 440 PRTYAF 445
```

RESULT 14

```
US-09-391-104-23
Sequence 23, Application US/09391104
Patent No. 6399371
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Falduto, Michael T.
APPLICANT: Magnuson, Scott R.
APPLICANT: Morgan, Douglas W.
TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
TITLE OF INVENTION: OF USING SAME
FILE REFERENCE: 6073.US.P1
CURRENT APPLICATION NUMBER: US/09/391,104
CURRENT FILING DATE: 1999-09-07
PRIOR APPLICATION NUMBER: US 08/814,394
PRIOR FILING DATE: 1997-03-11
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 469
TYPE: PRT
ORGANISM: Homo sapiens
US-09-391-104-23
```

```
Query Match 8.3%; Score 181.5; DB 4; Length 469;
Best Local Similarity 26.3%; Pred. No. 3.6e-10;
Matches 59; Conservative 26; Mismatches 82; Indels 57; Gaps 8;
```

```
QY 44 PVTTERCSDGMSFDA-TTLDNGTMLFFKGEFVWKSMDREL-----ISERKWNKPSPYD 98
DB 272 PQRPKACDSGLTFTDAITTL--RGEVWFFDRYFMRNRPYRVELANFISVFWQLNGSE 329
QY 99 AAFR-OGHNSVFLIKGDKVWVYPRPEKKEGYPLQDDEPFGIPSPDLAAVECHRGCEQAE 157
DB 330 AAYEFADRDVEVRFKGNKYWAVOGQVNLHGYPKDIYSF-GFP----- 371
QY 158 GVLFPQGHGRNGTGHNSTHGPEYRCSPHLVLSALSDNHGATTAFAFGTHVWRDL-- 215
DB 372 -----RTYKH-IDALSBENTGKTYFFVANKYRHYDEY 403
QY 216 -TSRDGMHSMPIAHQWPGSPAVDAAFSWEKLYLVQGTQVYVF 258
DB 404 KRSMDDPGYRPMIAHDFPGIGHKVDAVFMKDGPFFYFHTROYKF 447
```

RESULT 15

US-07-794-393-4

? Sequence 4, Application US/07794393
? Patent No. 5236844
? GENERAL INFORMATION:
? APPLICANT: CHAMON, PIERRE
? APPLICANT: BASSET, PAUL
? APPLICANT: BELLOCO, JEAN-PIERRE
? TITLE OF INVENTION: ANALYTICAL MARKERS FOR MALIGNANT BREAST
? TITLE OF INVENTION: CANCER
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Sterne, Kessler, Goldstein & Fox
? STREET: 1225 Connecticut Ave. NW Suite 300
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20036
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/794,393
? FILING DATE: 1991.12.1
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: GB 9025326.1
? FILING DATE: 21-NOV-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: GOLDSTEIN, JORGE A.
? REGISTRATION NUMBER: 29,021
? REFERENCE/DOCKET NUMBER: 1383.0040000
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 466-0800
? TELEFAX: (202) 833-8716
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 492 amino acids
? TYPE: AMINO ACID
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? IS-07-794-393-4

THIS PAGE BLANK (uspto)

QY 1 MARVLGAPALGLMSLCMSLAITATPLPTSAHGVAAGETKPPDVTETRCSDGMSFDATT 60

Db 1 MARVLGAPALGLMSLCMSLAITATPLPTSAHGVAAGETKPPDVTETRCSDGMSFDATT 60

QY 61 LDNGTMTLPFKKEPFYWKSHKMDRELISERKGNPSPVDAAPROGHSNVFLIKGDKWVYP 120

Db 61 LDNGTMTLPFKKEPFYWKSHKMDRELISERKGNPSPVDAAPROGHSNVFLIKGDKWVYP 120

QY 121 PEKKEGYPKLLIDEPFGI PSLPDAAVECHRGEOCABSVLPFO----- 163

DB 121 PEKKEKGYPLQDEFFPGIPSLDAVECHRGCEQAGVLPFGQDREMFMDLATGTWKEK 180
 QY 164 -----GCHGRN 169
 DB 181 SWPAVNCSSALMLGRYYCFOGNQFLRDPVGEVPPRYPRDVRDYFMPGPGRGHGRN 240
 QY 170 GTGHGNS 176
 DB 241 GTGHGNS 247

RESULT 2
 Q8KJ16 PRELIMINARY; PRT; 276 AA.
 ID Q8KJ16
 DT 01-OCT-2002 (TEMBLrel. 22, Created)
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Hemopexin (Fragment)
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN 1)
 RP SEQUENCE FROM N.A.
 RA Alam J., Smith A.;
 RT "Nucleotide sequence of the mouse haemopexin gene."
 RL Submitted (DEC-1990) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X56829; CAA40160.1; -
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 3.
 DR SMART; SM00120; HX; 3.
 FT NON TER 276
 SQ SEQUENCE 276 AA; 31047 MW; BD336C649BB2B767 CRC64;

Query Match 31.6%; Score 689.5; DB 11; Length 276;
 Best Local Similarity 53.0%; Pred. No. 7.8e-50;
 Matches 149; Conservative 12; Mismatches 41; Indels 79; Gaps 6;

QY 1 MAWVGLAPVALGMSLCSLAATPLPPTSAHGNVAEGE--TRDDPVYTRGSDGMSFDA 58
 DB 1 MARTAVANLTVLGLCWSLAVASPLP--TAHGRVAEVENGTKPSDVPBHCIDTWSFDA 58
 QY 59 TTLDGDTMLFFGGEFFWVKSHKMDRELISERWKNPSPVDAAFRQGNSTFLLKGDVWV 118
 DB 59 ATPDHNGTMLFFGGEFFWVKSHKMDRELISERWKNPSPVDAAFRQGNSTFLLKGDVWV 117
 QY 119 YPEKKEKGYPLQDEFFPGIPSLDAVECHRGCEQAGVLPFGQDREMFMDLATGTWKEK 165
 DB 118 YPEKKEKGYPLQDEFFPGIPSLDAVECHRGCEQAGVLPFGQDREMFMDLATGTWKEK 177
 QY 166 -----GH-- 167
 DB 178 ERSWPAVNCSTALRLWRLERYCFOGNKFLRFPNVTGEVPPRYPRDVRDYFVSCPGRGHGR 237
 QY 168 -RNGTGHGNSHTHGPXEMRCSPLVLVSLATSDNHGATYAPS 207
 DB 238 PRNGTGHGNSHTHGPXEMRCSPLVLVSLATSDNHGATYAPT 276

RESULT 3
 Q8J1P8 PRELIMINARY; PRT; 427 AA.
 ID Q8J1P8
 DT 01-OCT-2002 (TEMBLrel. 22, Created)
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Warm-temperature-acclimation-related-65 kDa-protein-like protein.
 GN WAP65-LIKE
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Belontiiformes; Adrianichthyidae; Oryziatidae; Oryzias.
 OX NCBI_TaxID=8090;
 RN 1)
 RP SEQUENCE FROM N.A.
 RA Hirayama M., Kobiyama A., Kinoshita S., Watabe S.;
 RT "Primary structural analysis of medaka Wap65 and Wap65-like protein
 RT and their gene expression in association with temperature acclimation
 RT and during ontogeny."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB075199; BAB97304.1; -
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 3.
 DR SMART; SM00120; HX; 7.
 SQ SEQUENCE 427 AA; 48000 MW; DEDB66A5BAD38EB CRC64;

Query Match 28.6%; Score 625; DB 13; Length 427;
 Best Local Similarity 33.0%; Pred. No. 3.5e-44;
 Matches 146; Conservative 66; Mismatches 136; Indels 94; Gaps 12;

QY 15 SLCSLAATPLPPTSAHGNVAEGE--TRDDPVYTRGSDGMSFDAATLDDNGTMLFFKGE 73
 DB 11 ALMLALTTRAAPLEDSAA--GDGSAALPDRCG--GIEFDATTPDDKQOTFFKCD 60
 QY 74 FVWKSHPMDRELISERWKNPSPVDAAFR-----QG-HNSVFLIKGDKVWVYPEKKEK 126
 DB 61 HWKKGFGGAGQSSQVFKLNGHVDAFRMHNHPENGDDHLYLPDDKVFSEFTLLE 120
 QY 127 GYPLQDEFFPGIPSLDAVECHRGCEQAGVLPFGQDREMFMDLATGTWKEK 164
 DB 121 GYPLQDEFFPGIPSLDAVECHRGCEQAGVLPFGQDREMFMDLATGTWKEK 180
 QY 165 -----HGR-----NGTGHGNSHTHGPXEMRCSPLVLVSLATSDNHGATYAPS 188
 DB 181 ACTSVFRLMEHYHCCFHHNPTFPQVYTGVTGYPKDAARFYRRCDFGHGERTTLKCS 240
 QY 189 HVLVSLATSDNHGATYAPSFGTGYWRLDTSRDGSHSPRIAHQPOGSAVDAAFSWEKLY 248
 DB 241 DPFMDALITDDGGRMYWFMFGSNVRLDTHRDGHARPIITSMKELNGDAVFSYNDRIY 300
 QY 249 LVQGTGVVYPLTKGTYLVSIGYPRKLEKEVGTPHGIILDSVDAAFCPGSSRLHIAGR 308
 DB 301 LKIGDQVYVYKAGAHFTLLIEGYPKTLKEBLN-----IEQVDAAPFCPORTAHIIQGRK 355
 QY 309 LWMLDLKSQAQATWELPWHKPKVGCALCMKSLGNSGANGPYLHGPVLYCYSDV 368
 DB 356 IYVINDLAATPREITLDAAPFGDIDAF-----SSDGITIKFGSNYHYDGP 403
 QY 369 EKLNAAKALPOPOVNTS-LIGC 389
 DB 404 MLVMSRIAPILPKVTSAMVGC 425

RESULT 4
 Q90WF7 PRELIMINARY; PRT; 439 AA.
 ID Q90WF7
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Warm-temperature-acclimation-related-65 kDa-protein precursor.
 GN WAP65.
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 NCBI_TaxID=7962;
 RN 1)
 RP SEQUENCE FROM N.A.
 RA Kinoshita S., Itoi S., Watabe S.;
 RT "cDNA Cloning and Characterization of the Warm-Temperature-
 RT Acclimation-Related-65 kDa-Protein."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB052623; BAB60809.1; -
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 5.
 DR SMART; SM00120; HX; 5.
 KM Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 439 POTENTIAL.
 SQ SEQUENCE 439 AA; 50045 MW; 24C9FB96629A9AD CRC64;

Query Match 21.3%; Score 464.5; DB 13; Length 439;
 Best Local Similarity 29.1%; Pred. No. 1.1e-30;
 Matches 136; Conservative 69; Mismatches 139; Indels 123; Gaps 18;

QY 15 SLCSLAIATPL-----PPTSAHGVABEGTKPDVTERCSGMSFDTATDDNGT 67
 Db 6 TLCLALSLFPAASDVADDPDTAGIKPELHHEAKD-----RCA-GMEFDIAVAVEGIP 59
 QY 68 LFFKGEFVWKSXKMDRELISERKNFP-----SPVDAFR-----QGHNSVFLIK 112
 Db 60 YFFKGDHLFKGFHGAELSNF-----TFPELDHNNLGHVDAAFRMHSESDSPHHDHQFFFL 116
 QY 113 GDKVWVYPEREKEXYPLDDEPRGIPSLDAVECHRGSCQAGVLFPG----- 164
 Db 117 DNKVFSTYKHLKLEKQDPDIDLPDHLDAAVECPKPDCTDVIFFKGDDELTFPM 176
 QY 165 -----HGR-----NGTGHNGSTHNGPRE-MRCS 187
 Db 177 KTKVDEKEFSMPCNCTGAFRYMEHYCFHGHQSKPRPVTDGQKPKETRDYFMR 235
 QY 188 PH-----LVLSALTSDNHGATYAFSGTHYRLDTSRDMHSPRIAHQW 230
 Db 236 PHFGQKSTEEHIEREQCSRVHLDAITSDDSIVAFRGYHF--VSITGDKFHSPTVESAF 293
 QY 231 PQGSVAUDAASWEKLYVGTQVYVFLTKGTYLVGCGYKLEKEVTHGIIIDSV 290
 Db 294 KELHSEVDAVFSYSGHLTMDKNEFYVYKVBSPHTLHGYEKPKLEVIG-----IEGRVD 348
 QY 291 AAFICPSSRLHIMAGRLMWLDLKSQAQ-----TWTELPMPHEKVGALCMESLCPN 345
 Db 349 AAFVCAIDHIAHIVIKGQTVYVDELKATPRAKKEITIQF-----KTIIAAMC-----GPK 399
 QY 346 SCANGPGLYLHGNPLCYSDVEKLANAKALPQPNVT-SLIGCTH 391
 Db 400 -----GVTAVIGNHYLYLDSPKIMMAKIMPEGRHSQGLFGGDH 439

RESULT 5

Q90WR3 PRELIMINARY; PRT; 261 AA.
 AC Q90WR3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hemopexin (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RX NCBI_TaxID=9031;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Weisbach L.; Plant P.; Grieninger G.;
 RT "Chicken hemopexin gene expression and sequence of a partial length
 RT CDA".
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M37319; AL29887.1; -
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 3.
 DR SMART; SM00120; HX; 3.
 FT NON_TER 1
 FT SEQUENCE 261 AA; 29385 MW; 8C2BBE7010F3C9A0 CRC64;

Query Match 20.9%; Score 456; DB 13; Length 261;
 Best Local Similarity 36.2%; Pred. No. 2.8e-30;
 Matches 108; Conservative 46; Mismatches 92; Indels 52; Gaps 11;

QY 96 PVDAAPFOGHNSVFLIKGDKWYVPEPKK--KQYPLDDEPRGIPSLDAVECHRG 153
 Db 10 PCDALRL-WLERYYCLOGTQFYRRPRHSEWLPYPRDLRDYF--ICP----- 55
 QY 154 CQAGVLFPGHGHGRTNGTGHNGSTHNGPREYMRCSPHLVLSALTSDNHGATYAFSGTHYMR 213
 Db 56 -----GGRH-----HGN-TSWGNAGDRCGE-PFOATISDSDGHYAFRGISFR 99
 QY 214 LDTSRDGMHSPRIAHQWPOGPSVADAFAFSWEKLYVGTQVYVFLT-KGQYTLVSGYPK 272
 Db 100 LDSWRDGMHAMPQHSNPGIAGVDAAFSMDKMYTLQSGSVSYVSGRGSHQVLEGYPR 159
 QY 273 RLEKEVTPHGIILDSVDAAPICGSSRLHIMAGRLMWLDLKSQAQATWELPHEKV 332
 Db 160 ALQELGVF-----KADAAFTCPGSAELYITGDRMQRVDTLTKSPHNADEPQPLPYDGV 213
 QY 333 DGLCMESLCPNCSANGPGLYLHGNPLCYSDVEKLANAKALPQPNVT-TSLIGC 389
 Db 214 DGAMC-----TADGTYLLRGDRYHRHMDVALLAAPPADPSTAVDLFHC 259

RESULT 6

Q90310 PRELIMINARY; PRT; 445 AA.
 AC Q90310;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAM temperature acclimation-related 65-kDa protein precursor.
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OC NCBI_TaxID=7957;
 RX NCBI_TaxID=7957;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle, and Hepatopancreas;
 RX MEDLINE=95340486; PubMed=7615502;
 RA Kikuchi K.; Yamashita M.; Watabe S.; Aida K.;
 RT "The warm temperature acclimation-related 65-kDa protein, Wap65, in
 RT goldfish and its gene expression".
 RT J. Biol. Chem. 270:17087-17092(1995).
 RN [2]
 RP SEQUENCE OF 1-149 FROM N.A.
 RC TISSUE=Hepatopancreas;
 RA Kikuchi K.; Watabe S.; Aida K.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D50437; BAA08928.1; -
 DR EMBL; D85428; BAA19835.1; -
 DR HSSP; P20058; IHXN.
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; hemopexin; 5.
 DR SMART; SM00120; HX; 4.
 KM Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 445 WARM TEMPERATURE ACCLIMATION-RELATED 65-
 FT KDA PROTEIN, WAP65.
 SQ SEQUENCE 445 AA; 50769 MW; 79C32EAB63BA8266 CRC64;

Query Match 20.5%; Score 447.5; DB 13; Length 445;
 Best Local Similarity 28.6%; Pred. No. 2.9e-29;
 Matches 132; Conservative 68; Mismatches 153; Indels 109; Gaps 16;

QY 15 SLCSLAIATPLPPTSAHGVABEGTKPDV-----TRCSDGMSFDTATDDNGT 68
 Db 8 TLCLALSLFPAASDVADDPDTAGIKPELHHEAKD-----RCA-GMEFDIAVAVEGIP 59
 QY 69 LFFKGEFVWKSXKMDRELISERKNFP-----SPVDAFR-----GHSVFLIKG 113

```

Db 67 FFKGDHLFKGFHDQALSTNE---TFPELDENHHLGHVDAAFRMSHSDSPAHHDHOFFLD 123
Qy 114 DKWVWVPRPEKKEGYPRLLODEPFPGIPSLDAAVECHRGCOAGVLFPGQ-----164
Db 124 TKVFSYKXKLEKQDPDISLELPGIPLDLDAVECPPLPCADDTIIFFGDEIYHLDK 163
Qy 165 -----HGH-----NGTGHGNSHTHGPXY-MRCS 168
Db 184 TKVDEKEFMSMNCGAFRMDHYVFRGHGSKPRPIGVEQAKPKETRIYFMRCP 242
Qy 189 H-----LVLSALTSNDNGATYAFSGTHYWRDLTSDRGMSWPIAHQWP 231
Db 243 HFQOKTDEHIEREQCSRVLDAITSDDGSAVAFRGHNP--LSITGDKFHSPTIESAFK 300
Qy 232 QGSSAUDAASFWEKLYLVGQTVYVFLTKGTYLVSGYPRKLEKVGTHGIIILDSYDA 291
Db 301 ELHSEVDVAFVSGHLYMIDNEVFYKVGEPHTEGYPKPEKVLG-----IEGVD 355
Qy 292 AFICPGSSRLHIMAGRLMWLDLKSQAQATWTELPWPH-EKVDGALCMEKSLGPNCSAN 350
Db 356 AFVCAADHIMAHVKGQTVYVDLKAITPRVYVKEGSIHLAKKIDVAMC-----GPK-----405
Qy 351 GPELYLHGNPLYCYSDVEKLNAAKALPDPQNT-SLIGCTH 391
Db 406 --GTVAVIGNHYVQFGSPMIMMAKIMPEQHRVSGGLFGCDH 445

RESULT 7
Q8UIP9 PRELIMINARY; PRT; 430 AA.
AC Q8UIP9;
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 23, Last annotation update)
DE Warm-temperature-acclimation-related-65 kDa-protein.
GN WAP65.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adiantichthyidae; Oryziinae; Oryzias.
OC NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HNT;
RA Hirayama M., Koshiyama A., Kinoshita S., Watabe S.;
RT "Primary structural analysis of medaka Wap65 and Wap65-like protein
RT and their gene expression in association with temperature acclimation
RT and during ontogeny."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB075198; BAB97303.1; -
DR InterPro: IPR000585; Hemopexin.
DR Pfam: PF00045; hemopexin; 2.
DR SMART: SM00120; HX; 4.
SQ SEQUENCE 430 AA; 49847 MW; A8BFC56B33AAEF15 CRC64;

Query Match 20.4%; Score 445.5; DB 13; Length 430;
Best Local Similarity 28.1%; Pred. No. 4e-29;
Matches 126; Conservative 67; Mismatches 155; Indels 101; Gaps 14;

Qy 16 LCWSLAIATPLPPTSAHGNVAEGETKPPDPTVTERCSGWSFDTATLTDNGTMLFFKGFV 75
Db 10 LCLALVLAW---ADHHEHRRKGA-----VRDCK-GIEMDAVAVNEBGIPTFFKEDHL 58
Qy 76 WKSHKMDRELISERWKNP-----SPVDAFR-----QGHNSVFLIKGDKXWVYPEK 123
Db 59 FKGFHGOAELSNKSFALDHHHLGHVDAAFRMHYEDDLNHDHMFPLDNKVAFAYYQHK 118
Qy 124 KEKGYKLLDDEFGIISPRLDAAVECHRGCOAGVLFPGQ-----164
Db 119 LEAGYPAISVEVFGIDPHDLDAAVECPKECEDSVLFFKKNELFHFYVKNKTVDERDFR 178
Qy 165 -----HGH-----RNGTGHGNSHTHGPXY-MRCS-----187

```

```

Db 179 SMPNCTSAFRMEHNYCFPHGKFSKFPKTKGSEVRGKYPKDAKFFMRCSKFPEDNDHEER 238
Qy 188 ---PHVLVSLATSDNGATYAFSGTHYWRDLTSDRGMSWPIAHQWPQSPAVDAFSWE 244
Db 239 ERCSRVHLDALITSDAGNIYAFRGHHYIRKDGNDTLKADTTESAKELHSEVDVAFSYN 298
Qy 245 EKLTVGQTVYVFLTKGTYLVSGYPRKLEKVGTHGIIILDSYDAFICGSSRLHIM 304
Db 299 SHLYMIDQLEVYRVRGEHHTLAGYPRKQVQALG-----IKGPIDAVFCODRHIAHI 353
Qy 305 AGRRLMWLDLKSQAQATWTELPWP-HEKVDGALCMEKSLGPNCSANGELYLHGNPLY 363
Db 354 KDRHWYDVMSATPRATNKRPISTILKYVDAMC-----GPK-----GVKVFGRNHYY 401
Qy 364 CYSDVEKLNAAKALPDPQNT-SLIGCTH 391
Db 402 HFESPKTFVAARALPEQHRIISLELFGCDH 430

RESULT 8
P79825 PRELIMINARY; PRT; 446 AA.
ID P79825;
AC P79825;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hemopexin-like protein (fragment). (Salmo gairdneri).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OC NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97142135; PubMed=8988369;
RA Mlot S., Duval J., Le Goff P.;
RT "Molecular cloning of a hemopexin-like cDNA from rainbow trout
RT liver."
RL DNA Seq. 6:311-318(1996).
DR EMBL: Z68112; CAA92147.1; -
DR HSSP: P20058; IHXN.
DR InterPro: IPR000585; Hemopexin.
DR Pfam: PF00045; hemopexin; 4.
DR SMART: SM00120; HX; 4.
FT NON_TER
SQ SEQUENCE 446 AA; 50454 MW; 13A8BAEA42B37727 CRC64;

Query Match 20.2%; Score 442; DB 13; Length 446;
Best Local Similarity 27.8%; Pred. No. 8.3e-29;
Matches 126; Conservative 64; Mismatches 168; Indels 96; Gaps 13;

Qy 16 LCWSLAIATPLPPTSAHGNVAEGETKPPD-----VTERCSGWSFDTATLTDNGTMLFF 70
Db 11 LCLVLAISHAHNHAHQGGEDEGHEGHGHLGLDRC-QGIEMDAVAVNEBGIPTFFK 69
Qy 71 KGFVFKSHKMDRELISERWKNP-----SPVDAFRQ-----GHSVFLIKGDKXWV 118
Db 70 KGGHVFGRPHKAELESNSFAELDDHHLGHVDAAFRLMHPDKPTVEHDHIFFLDTKVP 129
Qy 119 YPEKKEKGYKLLDDEFGIISPRLDAAVECHRGCOAGVLFPGQ-----164
Db 130 YHGHQLETFGPDQISVFGIIPDLDAVAVCRAPDCEBAVIFFKDELYHYNVTKKYV 169
Qy 165 -----HGH-----RNGTGHGNSHTHGPXY-MRCS-----187
Db 190 EKKEGMPNCTSAFRMEHNYCFPHGQFSKFPDKTGEVGRVPRKEARDYFMKSGKGYTT 249
Qy 188 -----PHVLVSLATSDNGATYAFSGTHYWRDLTSDRGMSWPIAHQWPQSPAVDA 239
Db 250 DHIERERCSRVLDAITSDAGNIYAFRGHHFLEQDAGNDVMAADTTIESDPKELHSEVDA 309

```



```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proctacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
RT "Immune-relevant (including acute phase) genes identified in the
RT livers of rainbow trout, Oncorhynchus mykiss, by means of suppression
RT subtractive hybridization.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF281339; AAC30014.1; -.
DR HSSP; P45452; IPEX.
DR InterPro; IPR000585; Hemopexin.
DR Pfam; PF00045; hemopexin; 1.
DR SMART; SMO0120; HX; 1.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 13379 MW; 4543D625B162B3BC CRC64;

Query Match 10.3%; Score 224; DB 13; Length 116;
Best Local Similarity 40.7%; Pred. No. 3e-11;
Matches 46; Conservative 20; Mismatches 35; Indels 12; Gaps 2;

Oy 64 NGTMLPFKGFVWVSKSHWDEBLISERKMFPS-----PVDAAR-----QGHNSVFLI 111
Db 1 NGHTFFPKGHLNMGVFGPAOVSSAPFKELDNYHGLHVDAAARMHNKEKREKGDHYFF 60
Oy 112 KGDVWVYPPPEKKEKGVKXLLDPEPPIPSPLDAVVECHDGEQOAEGLVFPQG 164
Db 61 LDDVVFYVYHNSLEEGYFKDQLDFGVFVSHVDAVCEPCGECNSDSVLFKKG 113

RESULT 11
O12945 PRELIMINARY; PRT; 453 AA.
AC O12945;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Vitronectin.
OS Gallus gallus (chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=White Leghorn; TISSUE=Retina;
RC MEDLINE=98090066; PubMed=9362471;
RA Martinez-Morales J.R., Barbae J.A., Marti E., Bovolenta P., Edgar D.,
RA Rodriguez-Tear A.;
RT "Vitronectin is expressed in the ventral region of the neural tube and
RT promotes the differentiation of motor neurons.";
RL Development 124:5139-5147(1997).
DR EMBL; Y11030; CA71914.1; -.
DR HSSP; P45452; IPEX.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF01033; Somatomedin_B; 1.
DR PRINTS; PRO0022; SOMATOMEDINB.
DR SMART; SMO0120; HX; 4.
DR SMART; SMO0201; SO; 1.
DR PROSITE; PS00024; HEMOPEXIN; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 1.
SQ SEQUENCE 453 AA; 51636 MW; 82B3BBD62B937C4 CRC64;

Query Match 9.2%; Score 200.5; DB 13; Length 453;
Best Local Similarity 25.8%; Pred. No. 1.6e-08;
Matches 91; Conservative 42; Mismatches 103; Indels 117; Gaps 20;

26 LPTPSAHGNVAEGTKP---DPVTERCSDGMSFDATTLDDNGTMLPFKGFVWVSKSHWD 82

```

```

Db      127 LHPTTTTTSDETRNPSLDDEPEELCSRK--PENAFTDLKNGSIYAFRGYFELDSS 185
Qy      83 -----RELISERKKNFSPVDAAFR-----OGHNSVFLIKGDKWVYPEPEKKEGYPLQLQ 133
Db      186 VRDGYPLTSDVW-GLEGPIIDAAFTTRINCQ--KTYLFFKSSQVWRFDGALDPEYPRDIS 242
Qy      134 DEFGPIPSIDAAVECHRGEGCOAGVLFPOGHGRNGTGNSTHHPERYWRCSPHLVLS 193
Db      243 EGEGLIPNDIDA-----FALPAH-----SYHNE-----267
Qy      194 ALTSDNHGATYAFSGTHYWRDLTSRDGMSWPIAHQMPQ-----GPSAV--DAAF---241
Db      268 -----RVYFFKGYKM-----SYDFAHQPTQAECEKSPSTVFNNHYAFMNRD 309
Qy      242 SMEEKLYVGTGVVYVLTGKGYTLVSGYPRKLEKE--VGTPIHILDSVDAAFICGSS 299
Db      310 SWED-----IFLSLFGSRMVGASSQRLISRDWGVPMQ--LDAAMAGRIYVSSR 356
Qy      300 RLHIMAGR-----LM-WLDKSGAQTWTLPWHEKYDGLC 337
Db      357 QPRRRSRHRKRYRNHRTNLGLMSWLN--SDSESTDSW-----LSSQC 403

```

RESULT 12

```

ID      081WC3      PRELIMINARY;      PRT;      634 AA.
AC      081WC3;
DT      01-MAR-2003 (TEMBLrel. 23, Created)
DT      01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT      01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE      Hypothetical protein (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RC      SEQUENCE FROM N.A.
RA      Straubeberg R.;
RL      Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC040507; AA040507.1; -.
KW      Hypothetical protein.
FT      NON TER
SQ      SEQUENCE 634 AA; 70086 MW; 0D4A78663BAA6AF CRC64;

```

Query Match 9.1%; Score 199; DB 4; Length 634;
 Best Local Similarity 25.0%; Pred. No. 3.2e-08;
 Matches 80; Conservative 31; Mismatches 123; Indels 86; Gaps 16;

```

Qy      5 LGAPVALGL-----WSLCWLSLAITPL-----PPTSAHGNVAEGETKPPDPVTERCS 51
Db      305 VGDBLRGLYEDKVRWOLYGVRESVPTAOPREPLLPBPDPNKSAPPKQDVPKRS 364
Qy      52 DGMSFDTLDDNGTMLFPKGEFVWKSMDRELISER-----WKNFP---SPVDAAF 101
Db      365 T--HFDVAVA-QINGEAFEPFKGKFWRLTR--DRHLVSLQPAQMRFWMLGLHLHDSVDAYV 420
Qy      102 -RQHSNVFLIKGDKWVYPEPEKKEGYPKLQDEFGIISPLDAAVECHRGEGCOAGVL 160
Db      421 ERTSDHIVFEFGKGRYVWFQDNVNEBEGYPRPVSD--FELPPGIDAA-----465
Qy      161 FFGGHGRNGTGNSTHHPERYWRCSPHLVLSLJTSDNHGATYAFSGTHYWRDLT--SRD 219
Db      466 FSWAHNDR-----TYFFDQQLYWRDHTRH 491
Qy      220 GHSWMPI-AHQWPGPSAVDAAFSWEK-KLYLVGTGVVYVLTGKGYTLVSGYPRKLEKE 277
Db      492 MDGYPFAGSPLMNGVSTLDDAMKMSDGAISFFRG-QEYWKVLDGELEVAQYDQSTARD 550
Qy      278 -----VGTPIHILDSVDA 292
Db      551 WLVCGDSDADGSVAAGVDAA 570

```

RESULT 13

```

ID      099745      PRELIMINARY;      PRT;      240 AA.
AC      099745;
DT      01-MAY-1997 (TEMBLrel. 03, Created)
DT      01-MAY-1997 (TEMBLrel. 03, Last sequence update)
DT      01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE      Metalloelastase (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RC      SEQUENCE FROM N.A.
RA      MEDLINE=94043200; PubMed=8226919;
RA      Shapiro S.D., Kobayashi D.K., Ley T.T.;
RT      "Cloning and characterization of a unique elastolytic
RT      metalloproteinase produced by human alveolar macrophages.";
RL      J. Biol. Chem. 266:23824-23829(1993).
RN      [2]
RC      SEQUENCE FROM N.A.
RA      MEDLINE=97187267; PubMed=9034720;
RA      Borden P., Heller R.A.;
RT      "Transcriptional control of matrix metalloproteinases and the tissue
RT      inhibitors of matrix metalloproteinases.";
RL      Crit. Rev. Eukaryot. Gene Expr. 7:159-178(1997).
RN      [3]
RC      SEQUENCE FROM N.A.
RA      Lin D., Duncan M., Allen E., Araujo R., Aparicio A., Chai A.,
RA      Chung E., Davis K., Pederspiel N., Hyman R., Kalman S., Komp C.,
RA      Kurdi O., Lashkari D., Lew H., Namath A., Oefner P., Roberts D.,
RA      Heller R., Davis R.W.;
RL      "Three Matrix Metalloproteinases on 81kb of Pl insert.";
RL      Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U78045; AAB36943.1; -.
DR      HSSP; P39900; LJIZ.
DR      InterPro; IPR000585; Hemopexin.
DR      Pfam; PF00045; hemopexin; 4.
DR      SMART; SM00120; HX; 4.
DR      PROSITE; PS00024; HEMOPEXIN; 1.
FT      NON TER
SQ      SEQUENCE 240 AA; 28406 MW; 3E620FAF17EB7CCC CRC64;

```

Query Match 9.0%; Score 196.5; DB 4; Length 240;
 Best Local Similarity 26.9%; Pred. No. 1.5e-08;
 Matches 65; Conservative 22; Mismatches 92; Indels 63; Gaps 10;

```

Qy      30 SAHGNVAEGETKPPD--VTERCSDGMSFDTLDDNGTMLFPKGEFVW-----KSHKWRD 83
Db      30 SLYGDKREKQRLPNPNKSPALCDPRLSFDAYVTYGN-KIFFKDFPVLKYSERKTSV 88
Qy      84 ELISERKKNFSPVDAAFR--OGHNSVFLIKGDKWVYPEPEKKEGYPKLQD--EFGIIS 141
Db      89 NLISLMPRTLPSGIGBAVEIEANQVLFEPKDXWYLSMLRPPNPKSIHSFGFPNPKY 148
Qy      142 PLDAVECHRGEGCOAGVLFPOGHGRNGTGNSTHHPERYWRCSPHLVLSLJTSDNHG 201
Db      149 KIDAAV-----FNPFRFR-----161
Qy      202 ATYAFSGTHYWRDLTSDR-----GHSWPIAHQWPGPSAVDAAFSWEK-LYLVGTGVY 256
Db      162 -TYFPVNDQYWRDERKQMMDDGYPRLITKNPGIGPK-IDAVFYKANKYYVFFQSSNGP 219
Qy      257 VF 258
Db      220 EY 221

```

RESULT 14

```

ID      09TV75      PRELIMINARY;      PRT;      464 AA.
Q9TV75

```

```

RA Sorensen K.C., Mitchell B.E.;
RT "Comparison of feline scromelysin-1 (MMP3) cDNA with other known MMP3
RT cDNA sequences."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033627; AAK64605.1; -.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Matrxin.
DR InterPro; IPR006026; Nzn_Mtpeptidase.
DR InterPro; IPR006025; Zn_Mtpeptidase.
DR Pfam; PF00045; hemopexin; 3.
DR Pfam; PF00413; peptidase M10; 1.
DR Pfam; PF03933; peptidase M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 3.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 393 AA; 44093 MW; 148C78EAA6BD4A95 CRC64;

Query Match 8.7%; Score 189.5; DB 6; Length 393;
Best Local Similarity 26.4%; Pred. No. 1,1e-07;
Matches 63; Conservative 23; Mismatches 94; Indels 59; Gaps 8

OY 25 PLPTSAHGVABGET-KRDPDYTERCSDGWSPDATTLDNGTMLFFKGEFVWKS HKMDR 83
DB 201 PPMSSPDGPAGPTESVPEPCTPATCDPALSPDAVS-SLNGEILFFDRLHMRKSLRTR 259
OY 84 E-----LISERKKNFSPVDAFRO-GHNSVFLIKGDKVWVYPRPKKEKGYPKLQD-EFP 137
DB 260 EPGFYLMGCSFMPSLPSGLDAAVEETSKDIVPIFKGNQPMARKGTGEVQAGYPRKGIITLGF 319
OY 138 GIPSPDLAAVECHRGECQAEGLVLFQGHGHRNGTGHGNSHTHGRPEYMRCSPLVLVSALT 197
DB 320 PTVKKIDAAV-----FD 331
OY 198 DNHGATYAFSGTHYWRLDTSRDGMH-SWP--IAHQWPGSPAVDAAFSWEKLYLVQGT 253
DB 332 KEKKKYTFVGVGDKWRFDEKKGSMGPEPFKQIAEDPFGVDSKVDAAFEAFGYYFFNGS 390

Search completed: December 16, 2003, 06:52:11
Job time : 179 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 23:41:46 ; Search time 25 Seconds
(without alignments)
735.498 Million cell updates/sec

Title: US-09-900-448-2
Perfect score: 2185
Sequence: 1 MARVLGAPVALGLWSLCSL.....NAKALPQPQVNTSLGCTH 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2139.5	97.9	462	1 HEMO_HUMAN	P02790 homo sapien
2	1700.5	77.8	460	1 HEMO_RABIT	P20058 oxyctolagus
3	1556.5	71.2	460	1 HEMO_RAT	P20059 rattus norv
4	1545.5	70.7	460	1 HEMO_MOUSE	O91x72 mus musculu
5	1544	70.7	459	1 HEMO_PIG	P58828 sus scrofa
6	203	9.3	578	1 MM17_MOUSE	Q97083 mus musculu
7	202.5	9.3	470	1 MM12_HUMAN	P33900 homo sapien
8	199	9.1	606	1 MM17_HUMAN	O91129 homo sapien
9	196	9.0	464	1 MM12_RABIT	P73227 oxyctolagus
10	195.5	8.9	476	1 MM10_MOUSE	O55123 mus musculu
11	195.5	8.9	477	1 MM03_HORSE	Q28397 equus caball
12	189.5	8.7	476	1 MM10_HUMAN	P09238 homo sapien
13	187	8.6	476	1 MM10_RAT	P07152 rattus norv
14	184.5	8.4	468	1 MM01_RABIT	P13943 oxyctolagus
15	184.5	8.4	492	1 MM11_MOUSE	Q02853 mus musculu
16	183.5	8.4	467	1 MM08_HUMAN	P02894 homo sapien
17	181.5	8.3	469	1 MM01_HUMAN	P03956 homo sapien
18	179.5	8.2	488	1 MM11_HUMAN	P24347 homo sapien
19	177.5	8.1	471	1 MM13_RABIT	O62806 oxyctolagus
20	177	8.1	475	1 MM03_RAT	P03957 rattus norv
21	176	8.1	477	1 MM03_HUMAN	P08254 homo sapien
22	176	8.1	471	1 MM13_HUMAN	P45452 homo sapien
23	176	8.1	472	1 MM13_HORSE	O19927 equus caball
24	174.5	8.0	469	1 MM01_HORSE	Q95825 equus caball
25	173	7.9	477	1 MM03_MOUSE	P28862 mus musculu
26	172	7.9	669	1 MM15_HUMAN	P51511 homo sapien
27	169.5	7.8	477	1 MM11_XENLA	O11005 xenopus lae
28	169.5	7.8	478	1 VTNC_HUMAN	P04004 homo sapien
29	167.5	7.7	508	1 MM19_HUMAN	O99542 homo sapien
30	166.5	7.6	466	1 MM13_RAT	P23097 rattus norv
31	166	7.6	469	1 MM01_PIG	P23097 rattus norv
32	166	7.6	478	1 MM03_RABIT	P28863 oxyctolagus
33	165.5	7.6	478	1 VTNC_MOUSE	P29788 mus musculu

34	165	7.6	657	1 MM15_MOUSE	O54732 mus musculu
35	163	7.5	618	1 MM24_MOUSE	O97082 mus musculu
36	163	7.5	618	1 MM24_RAT	O99066 rattus norv
37	162.5	7.4	645	1 MM24_HUMAN	O94572 homo sapien
38	161.5	7.4	465	1 MM08_MOUSE	O70138 mus musculu
39	159.5	7.3	466	1 MM08_RAT	O88766 rattus norv
40	158	7.2	459	1 VTNC_PIG	P48819 sus scrofa
41	155.5	7.1	467	1 MM18_XENLA	O13065 xenopus lae
42	154.5	7.1	472	1 MM13_MOUSE	P33435 mus musculu
43	152	7.0	607	1 MM16_HUMAN	P55152 homo sapien
44	151	6.9	469	1 MM01_BOVIN	P28053 bos taurus
45	151	6.9	607	1 MM16_MOUSE	O9w10 mus musculu

ALIGNMENTS

```

RESULT 1
ID      HEMO_HUMAN      STANDARD;      PRT;      462 AA.
AC      P02750;
DT      21-JUL-1986 (Rel. 01, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hemopexin precursor (Beta-1B-glycoprotein).
GN      HPX.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      MEDLINE=88316972; PubMed=2842511;
RA      Altruda F., Poli V., Restagno G., Silengo L.;
RT      "Structure of the human hemopexin gene and evidence for
RT      intron-mediated evolution."
RT      J. Mol. Evol. 27:102-108(1988).
RN      [2]
RP      SEQUENCE OF 2-462 FROM N.A.
RX      MEDLINE=89122012; PubMed=3220477;
RA      Law M.L., Cai G.Y., Hartz J.A., Jones C., Kao F.T.;
RT      "The hemopexin gene maps to the same location as the beta-globin gene
RT      cluster on human chromosome 11."
RT      Genomics 3:48-52(1988).
RN      [3]
RP      SEQUENCE OF 22-462 FROM N.A.
RX      MEDLINE=85242073; PubMed=2989777;
RA      Altruda F., Poli V., Restagno G., Argos P., Cortese R., Silengo L.;
RT      "The primary structure of human hemopexin deduced from cDNA sequence:
RT      evidence for internal, repeating homology."
RT      Nucleic Acids Res. 13:3841-3859(1985).
RN      [4]
RP      SEQUENCE OF ACTIVE PROTEIN.
RX      MEDLINE=85113173; PubMed=385550;
RA      Takahashi N., Takahashi Y., Putnam F.W.;
RT      "Complete amino acid sequence of human hemopexin, the heme-binding
RT      protein of serum."
RT      Proc. Natl. Acad. Sci. U.S.A. 82:73-77(1985).
RN      [5]
RP      SEQUENCE OF 24-255.
RX      MEDLINE=85076955; PubMed=6510521;
RA      Frantlova V., Borvak J., Klub I., Moravsek L.;
RT      "Amino acid sequence of the N-terminal region of human hemopexin."
RT      FEBS Lett. 178:213-216(1984).
RN      [6]
RP      PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RX      MEDLINE=84193947; PubMed=6371807;
RA      Takahashi N., Takahashi Y., Putnam F.W.;
RT      "Structure of human hemopexin: O-glycosyl and N-glycosyl sites and
RT      unusual clustering of tryptophan residues."
RT      Proc. Natl. Acad. Sci. U.S.A. 81:2021-2025(1984).
CC      -I- FUNCTION: BINDS HEME AND TRANSPORTS IT TO THE LIVER FOR BREAKDOWN
AND IRON RECOVERY, AFTER WHICH THE FREE HEMOPEXIN RETURNS TO THE

```

CC CIRCULATION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
 CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M36803; AA58678.1; -.
 DR EMBL: M36796; AA58678.1; JOINED.
 DR EMBL: M36799; AA58678.1; JOINED.
 DR EMBL: M36800; AA58678.1; JOINED.
 DR EMBL: M36801; AA58678.1; JOINED.
 DR EMBL: M36802; AA58678.1; JOINED.
 DR EMBL: J03048; AA52704.1; -.
 DR EMBL: X02537; CAA26382.1; ALT_INIT.
 DR PIR: I56456; OOHU.
 DR HSSP: P20058; IHXN.
 DR SWISS-2DPAGE: P02790; HUMAN.
 DR Sienna-2DPAGE: P02790; -.
 DR GeneW: HGNC:5171; HPX.
 DR MIM: 142290; -.
 DR GO: GO:0005615; C:extracellular space; TAS.
 DR GO: GO:0005488; F:binding activity; TAS.
 DR GO: GO:0015232; F:heme transporter activity; TAS.
 DR GO: GO:0005211; F:plasma glycoprotein; TAS.
 DR GO: GO:0015886; P:heme transport; TAS.
 DR GO: GO:0006879; P:iron ion homeostasis; TAS.
 DR InterPro: IPR005058; Hemopexin.
 DR Pfam: PF00045; hemopexin; 5.
 DR SMART: SM00120; HX; 5.
 DR PROSITE: PS00024; HEMOPEXIN; 2.
 DR GlycoProtein: Heme; Plasma; Repeat; Transport; Signal.
 DR KW Glycoprotein; Heme; Plasma; Repeat; Transport; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 462 HEMOPEXIN-LIKE 1.
 FT DOMAIN 56 93 HEMOPEXIN-LIKE 2.
 FT DOMAIN 97 141 HEMOPEXIN-LIKE 3.
 FT DOMAIN 188 231 HEMOPEXIN-LIKE 4.
 FT DOMAIN 263 306 HEMOPEXIN-LIKE 5.
 FT DOMAIN 308 351 HEMOPEXIN-LIKE 5.
 FT METAL 79 79 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL 150 150 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT DISULFID 50 231
 FT DISULFID 149 154
 FT DISULFID 188 200
 FT DISULFID 257 460
 FT DISULFID 366 408
 FT DISULFID 418 435
 FT CARBOHYD 24 24 O-LINKED.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 462 AA; 51676 MW; 054B44D603763B8 CRC64;
 Query Match 97.9%; Score 2139.5; DB 1; Length 462;
 Best Local Similarity 84.6%; Pred. No 5,8e-176;
 Matches 391; Conservative 0; Mismatches 0; Indels 71; Gaps 1;

QY 121 PEKKEKGYPLQDEFPPIPSPLDAAVECHGECAEVLFPQ----- 163
 DB 121 PEKKEKGYPLQDEFPPIPSPLDAAVECHGECAEVLFPQGDREWFMDLATGTMKER 180
 QY 164 -----GHGHRN 169
 DB 181 SWPAVNCSSALRWLGRYYCFQGNQELRPDPVGEVPPRYPRDVRDYFMPCCPGRGHGRN 240
 QY 170 GTGHSNSTHGPPEYMCSPHLVLSALTSNDHGATVAFSGTHYWRDLTSDRGHMSPIAHQ 229
 DB 241 GTGHSNSTHGPPEYMCSPHLVLSALTSNDHGATVAFSGTHYWRDLTSDRGHMSPIAHQ 300
 QY 230 WPGPSAVDAFAFSEWEKYLVOCTQYVFLTKGTYLVSGPYRLEKEVGTPIHILDSV 289
 DB 301 WPGPSAVDAFAFSEWEKYLVOCTQYVFLTKGTYLVSGPYRLEKEVGTPIHILDSV 360
 QY 290 DAAFTCPGSSRLHMGRLMWLDLKSQAQATTELPMWHEKVDALCKEKSIGPNCSA 349
 DB 361 DAAFTCPGSSRLHMGRLMWLDLKSQAQATTELPMWHEKVDALCKEKSIGPNCSA 420
 QY 350 NGPGLYLHGPMLCYSDVEKLNAAKALPOPNVTSILGCTH 391
 DB 421 NGPGLYLHGPMLCYSDVEKLNAAKALPOPNVTSILGCTH 462

RESULT 2

HEMO_RABBIT
 ID HEMO_RABBIT STANDARD; PRT; 460 AA.
 AC P20058;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hemopexin precursor.
 GN HPX.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_Taxid=9986;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND HEME-BINDING SITES.
 RC TISSUE=Liver;
 RX MEDLINE=93203213; PubMed=7681064;
 RA Morgan W.T., Master P., Tatum F., Kao S.-M., Alm J., Smith A.;
 RT "Identification of the histidine residues of hemopexin that
 coordinate with heme-iron and of a receptor-binding region.";
 RL J. Biol. Chem. 268:6256-6262(1993).
 RN [2]
 RP SEQUENCE OF 26-53.
 RX MEDLINE=88339942; PubMed=3421961;
 RA Wellner D., Cheng K.C., Mueller-Eberhard U.;
 RT "N-terminal amino acid sequences of the hemopexins from chicken, rat
 and rabbit.";
 RL Biochem. Biophys. Res. Commun. 155:622-625(1988).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 242-460.
 RX MEDLINE=96173004; PubMed=8590016;
 RA Faber H.R., Groen C.R., Baker H.M., Morgan W.T., Smith A., Baker E.N.;
 RT "1.8-A crystal structure of the C-terminal domain of rabbit serum
 haemopexin.";
 RL Structure 3:551-559(1995).
 CC -1- FUNCTION: BINDS HEME AND TRANSPORTS IT TO THE LIVER FOR BREAKDOWN
 CC AND IRON RECOVERY, AFTER WHICH THE FREE HEMOPEXIN RETURNS TO THE
 CC CIRCULATION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
 CC -1- SIMILARITY: Contains 5 hemopexin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

Db 389 VDAVERTSDHKIVFPGKDRYVWFVKDNNVEBGYPREVSD-FSLPFGIDA----- 437
 QY 156 AEGVLFPQGHGHRNGTHGNGSTHHPYMCSPHLVLSALTSNKGATYAFSGTHYRLD 215
 Db 438 ----VFSWANDR-----TYFFKQOLYWRVD 459
 QY 216 --TSR--DGWHS-WPIAHQWPGPSAVDAFAFWEE-KLYLVQSTQYVFLTKGTYLVSG 269
 Db 460 DHTRRMDPGYPAGCPFL--WRGVPSMLDDMRKWDGASFFRG-QEYWKVLDELBAAPG 515
 QY 270 YPKRLEKE--VGTPHGIIILSDVDAFICPGSSRLHIMAGRRLMWLDLKSQAQATWTELP 326
 Db 516 YPQSTADMLVCGEP-----LADAEDVGRPGQ-----GR-----SGAQ----- 548
 QY 327 WPHKXVDG-ALCMEKSLGPNCSANGPGVLYIHGPNLYCYSDVEKLNAAKALPQPNVTS 385
 Db 549 ----DGLAVCS-----CTSDAHLR-----ALPSLLLLTP 573
 QY 386 LL 387
 Db 574 LL 575
 RESULT 5
 AAY8405
 ID AAY8405 standard; Protein; 587 AA.
 AC AAY8405;
 XX 28-JUL-2000 (first entry)
 DE Mouse MT4-MMP (2) matrix metalloproteinase amino acid sequence.
 XX
 KM Matrix metalloproteinase; MT4-MMP; treatment; arthritis deformans;
 KM chronic rheumatoid arthritis; asthma; autoimmune disease; hepatitis;
 KM atrophic dermatitis; psoriasis; contact dermatitis; hair loss; nephritis;
 KM ischaemic disease; pancreatitis; arteriosclerosis; leukemia; tumour;
 KM wound; corneal ulcer; tissue damage; inflammation; cerebral apoplexia;
 KM Alzheimer's disease; dementia; multiple sclerosis; Parkinson's disease;
 KM gene therapy.
 KM
 OS Mus sp.
 XX
 PN W0200018900-A1.
 PD 06-APR-2000.
 XX
 PF 29-SEP-1999; 99WO-JP05349.
 XX
 PR 29-SEP-1998; 98JP-0276258.
 PR 29-SEP-1998; 98JP-0291505.
 XX
 PA (SEIKI/) SEIKI M.
 XX
 PI Seiki M;
 DR WPI: 2000-293141/25.
 DR N-PSDB; AAA13378.
 XX
 PT Physiologically-active transmembrane matrix metalloproteinase
 PT polypeptide, useful in screening inhibitors and activators for treating
 PT e.g. arthritis deformans, asthma and cancers
 XX
 PS Claim 11; Page 59-66; 113pp; Japanese.
 XX
 CC This sequence represents the murine matrix metalloproteinase (MT4-MMP 2)
 CC amino acid sequence. The invention relates to a mammalian transmembrane
 CC MMP polypeptide, its variants and the DNA sequences encoding them. Also
 CC included in the invention is a vector comprising the MMP DNA molecule,
 CC and cells transformed using the vector. The protein, DNA and antibodies
 CC directed against the polypeptide are useful in screening inhibitors and
 CC activators for use in treating arthritis deformans, chronic rheumatoid
 CC arthritis, asthma, autoimmune diseases, atrophic dermatitis, psoriasis,

CC contact dermatitis, hair loss, ischaemic diseases, immune reaction
 CC accompanying organ transplant, hepatitis, nephritis, pancreatitis,
 CC arteriosclerosis, leukaemia, malignant tumours, wounds, corneal ulcers,
 CC tissue damage or inflammation accompanying leukocytic infiltration,
 CC together with brain disorders during cerebral apoplexia, Alzheimer's
 CC disease, dementia, multiple sclerosis, Parkinson's disease or brain
 CC tumours. The DNA can also be incorporated into a suitable vector for use
 CC in gene therapy.
 XX
 SQ Sequence 587 AA;
 Query Match 9.3%; Score 204; DB 21; Length 587;
 Best Local Similarity 24.4%; Pred. No. 3.9e-10;
 Matches 103; Conservative 34; Mismatches 127; Indels 158; Gaps 25;
 QY 5 LGAPVALGL-----WSLCSLAIATPL-----PPPSANGNVAEGTKDDPY 46
 Db 273 VGDPRVYGGPYEDRVRVMOLYGVRESVSPQAQDTPPEPBPPLPPPPNNRSSTPQKDV 332
 QY 47 TERCSDGWSFDATTLDDNGTMLPFKGFVWKSHKMDRELISER-----WKNFP---SP 96
 Db 333 PNRCTA--HFDAVA-QIRGEAFFKGFWRRLTR-DRLVSLQPAQMRFRGLPLHLDS 388
 QY 97 VDAAF-ROGHNSVFLIKGDKVWVYPPKXKXGYKLLQDEFPGIPSPPLDAAYVCHRGECQ 155
 Db 389 VDAVERTSDHKIVFPGKDRYVWFVKDNNVEBGYPREVSD-FSLPFGIDA----- 437
 QY 156 AEGVLFPQGHGHRNGTHGNGSTHHPYMCSPHLVLSALTSNKGATYAFSGTHYRLD 215
 Db 438 ----VFSWANDR-----TYFFKQOLYWRVD 459
 QY 216 --TSR--DGWHS-WPIAHQWPGPSAVDAFAFWEE-KLYLVQSTQYVFLTKGTYLVSG 269
 Db 460 DHTRRMDPGYPAGCPFL--WRGVPSMLDDMRKWDGASFFRG-QEYWKVLDELBAAPG 515
 QY 270 YPKRLEKE--VGTPHGIIILSDVDAFICPGSSRLHIMAGRRLMWLDLKSQAQATWTELP 326
 Db 516 YPQSTADMLVCGEP-----LADAEDVGRPGQ-----GR-----SGAQ----- 548
 QY 327 WPHKXVDG-ALCMEKSLGPNCSANGPGVLYIHGPNLYCYSDVEKLNAAKALPQPNVTS 385
 Db 549 ----DGLAVCS-----CTSDAHLR-----ALPSLLLLTP 573
 QY 386 LL 387
 Db 574 LL 575
 RESULT 6
 AAE10419
 ID AAE10419 standard; Protein; 470 AA.
 AC AAE10419;
 XX
 DT 10-DEC-2001 (first entry)
 DE Human matrix metalloproteinase-12 (MMP-12) protein.
 XX
 KM Human; matrix metalloproteinase; MMP-12; hair growth; antisense therapy;
 KM endopeptidase; skin cell; breast cancer; hair follicle; chromosome 11q22.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Peptide
 FT 1..16 Location/Qualifiers
 FT /label= Signal_peptide
 FT 17..470
 FT /label= Mature_MMP_12_protein
 FT 90..96
 FT /label= Cysteine_switch_domain
 FT 161..185
 FT Domain
 FT /note= "Zinc and calcium binding domain"
 XX

OS Homo sapiens.
 XX WO200190366-A2.
 XX 29-NOV-2001.
 XX 24-MAY-2001; 2001WO-US17076.
 XX 24-MAY-2000; 2000US-206690P.
 XX (CURA-) CURAGEN CORP.
 XX Leach MD, Shinkets RA;
 XX WPI; 2002-106200/14.
 XX N-PSDB; ABN75603.
 DR Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation
 XX Claim 10; Page 541-542; 2508pp; English.
 XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN75587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening individuals for ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, hematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, tumour inhibition activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antifibrotic activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 XX
 XX Sequence 87 AA;
 SQ
 Query Match 11.8%; Score 258; DB 23; Length 87;
 Best Local Similarity 86.5%; Pred. No. 2e-16; Mismatches 0; Gaps 0;
 Matches 45; Conservative 3; Indels 0; Indels 0; Gaps 0;
 QY 114 DKWVVPPEKKEKGYPLQLQDEFFGIPSPDAAVECHRGCGOAGVLFQGH 165
 DB 5 DKWVVPPEKKEKGYPLQLQDEFFGIPSPDAAVECHRGCGOAGVLFQGH 56
 RESULT 4
 AAY90502
 ID AAY90502 standard; Protein; 587 AA.

XX AAY90502;
 AC 15-AUG-2000 (first entry)
 XX
 DT Murine transmembrane matrix metalloproteinase MT4-MMP(2).
 DE
 XX Matrix metalloproteinase; MT4-MMP(2); transmembrane; mouse; antibody;
 XX drug screening; diagnosis; arthrosis deformans; rheumatoid arthritis;
 KW asthma; autoimmune disease; atopic dermatitis; skin disorder;
 KW ischaemic disorder; arteriosclerosis; cancer; tissue damage;
 KW inflammatory disorder; neurological disorder.
 XX
 XX Mus sp.
 OS
 XX WO200018805-A1.
 XX 06-APR-2000.
 XX 29-SEP-1999; 99MO-JP05350.
 XX 29-SEP-1998; 98JP-0291501.
 XX 29-SEP-1998; 98JP-0291503.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Hanai N, Furuya A;
 XX WPI; 2000-293115/25.
 XX N-PSDB; AAA14339.
 XX Claim 1; Page 86-90; 152pp; Japanese.
 XX This sequence represents a novel murine transmembrane matrix
 CC metalloproteinase MT4-MMP(2). MT4-MMP(2) has physiological activity
 CC different to that of MT4-MMP. The invention relates to antibodies
 CC against both murine and human MT4-MMP(2) (AAY90502 and AAY90503) and to
 CC antibodies against a novel murine and human transmembrane matrix
 CC metalloproteinase MT5-MMP (AAY90504 and AAY90505). The antibodies are
 CC useful for the diagnosis and screening of inhibitors and activators
 CC useful for treating or preventing a wide variety of medical conditions,
 CC such as arthrosis deformans, rheumatoid arthritis, asthma, autoimmune
 CC diseases and atopic dermatitis. The anti-MT5-MMP antibodies may
 CC additionally be used to diagnose and screen therapeutic or prophylactic
 CC agents for conditions affecting the brain, such as cerebral stroke and
 CC Alzheimer's disease. The antibodies of the invention may also be used for
 CC the prevention, diagnosis and treatment of psoriasis, contact
 CC dermatitis, hair loss, ischaemic diseases, immune reaction accompanying
 CC organ transplant, hepatitis, nephritis, pancreatitis, arteriosclerosis,
 CC leukemia, malignant tumours, wounds, corneal ulcers, tissue damage or
 CC inflammation accompanying leukocytic infiltration, dementia, multiple
 CC sclerosis, Parkinson's disease or brain tumour.
 XX
 XX Sequence 587 AA;
 SQ
 Query Match 9.3%; Score 204; DB 21; Length 587;
 Best Local Similarity 24.4%; Pred. No. 3.9e-10;
 Matches 103; Conservative 34; Mismatches 127; Indels 158; Gaps 25;
 QY 5 LGAPVALGI-----WSLCWSLAATPL-----PPTSAHGNAVAGETKPDV 46
 DB 273 VSDPVYVGGPYEDRVRVWVQYGVRESVSPTAQLDTPPEPEPLPPEPNNRSTPQKV 332
 QY 47 TERCSGNSGSPDATTLDNGTMLPFKGEFVWKSXKMDRELISR-----WKNFP---SP 96
 DB 333 PRRCTA--HFDVA--QIRGEAFPFKGYKFWRLTR--DRHLVSLQPAQWHRFWGLPLHDS 388
 QY 97 VDAAF--RQGNNSVFLTKGDKWVVPPEKKEKGYPLQLQDEFFGIPSPDAAVECHRGCGO 155

XX WPI: 2001-476164/51.
 DR N-PSDB; AAB98592.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 XX antibodies and research use -
 XX
 PS Claim 20; Page 1017-1018; 1275pp; English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.
 XX
 SQ Sequence 462 AA;
 Query Match 97.9%; Score 2139.5; DB 22; Length 462;
 Best Local Similarity 84.6%; Pred. No. 8.9e-193;
 Matches 391; Conservative 0; Mismatches 0; Indels 71; Gaps 1;
 QY 1 MARVLGAPVALGLMSLCWSLAIATPLPPTSAHGNVAGETKPPDPTERTCSDGMSFDATT 60
 DB 1 MARVLGAPVALGLMSLCWSLAIATPLPPTSAHGNVAGETKPPDPTERTCSDGMSFDATT 60
 QY 61 LDDNGTMLPFKGEFVWKSMDRELISERWKNFPPDAFROGHSNVPFLIKGDKWVVP 120
 DB 61 LDDNGTMLPFKGEFVWKSMDRELISERWKNFPPDAFROGHSNVPFLIKGDKWVVP 120
 QY 121 PERKEKGYPLTLDPEFGIPSPLDAAVECHRGECQAEVLPFO----- 163
 DB 121 PERKEKGYPLTLDPEFGIPSPLDAAVECHRGECQAEVLPFO----- 163
 QY 164 -----GHGHRN 169
 DB 181 SWPVGNCSSALMKWLGHYTFQGNQFLRPPVGEVPPRYPRDVRDYFMPGGRGHNRN 240
 QY 170 GTGHGNSHHGPEYMRCSPLVLSALTSDNHGATYAFSGTHYRLDTSRDGMSWPIAHQ 229
 DB 241 GTGHGNSHHGPEYMRCSPLVLSALTSDNHGATYAFSGTHYRLDTSRDGMSWPIAHQ 300
 QY 230 WPGGSAVDAAFSWEEKLYVQGTQVYVFLTKGTYTVSGYPKLEKEVGTPIGIIIDSV 289
 DB 301 WPGGSAVDAAFSWEEKLYVQGTQVYVFLTKGTYTVSGYPKLEKEVGTPIGIIIDSV 360
 QY 290 DAAPFCGSSRLHIMAGRLMWLIDLSGAQATWTELPMPHEKVDGALCMESKISGPNCSA 349
 DB 361 DAAPFCGSSRLHIMAGRLMWLIDLSGAQATWTELPMPHEKVDGALCMESKISGPNCSA 420
 QY 350 NGPGLYLHGPNTLYCYSDEKLNAAKALPOPONTSLIGCTH 391
 DB 421 NGPGLYLHGPNTLYCYSDEKLNAAKALPOPONTSLIGCTH 462
 RESULT 2
 AAG00304
 ID AAG00304 standard; Protein; 83 AA.
 XX
 AC AAG00304;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 4385.
 XX
 KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.

XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-020610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI: 2000-500381/45.
 XX
 DR N-PSDB; AAC00310.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 4385; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 83 AA;
 Query Match 20.9%; Score 456; DB 21; Length 83;
 Best Local Similarity 98.8%; Pred. No. 3.9e-35;
 Matches 82; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MARVLGAPVALGLMSLCWSLAIATPLPPTSAHGNVAGETKPPDPTERTCSDGMSFDATT 60
 DB 1 MARVLGAPVALGLMSLCWSLAIATPLPPTSAHGNVAGETKPPDPTERTCSDGMSFDATT 60
 QY 61 LDDNGTMLPFKGEFVWKSMDRELISERWKNFPPDAFROGHSNVPFLIKGDKWVVP 120
 DB 61 LDDNGTMLPFKGEFVWKSMDRELISERWKNFPPDAFROGHSNVPFLIKGDKWVVP 120
 RESULT 3
 ABP31577
 ID ABP31577 standard; Protein; 87 AA.
 XX
 AC ABP31577;
 XX
 DT 09-JUL-2002 (first entry)
 XX
 DE Human glycoprotein-like ORF550 protein, SEQ ID NO:1100.
 XX
 KM Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;
 KM immune modulation; haematopoiesis regulation; tissue growth;
 KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KM thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KM behaviour; cancer; proliferative disorder; neurological disorder;
 KM cardiovascular disease; immune system disorder; organ transplantation;
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KM hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KM vasotropic; antiporatic; antidiabetic; cytostatic; neurotropic;
 KM neuroprotective; antilatherosclerotic; anticoncugant; thrombolytic;
 KM cardiant; hypotensive; antichyloid; antinflammatory; immunomodulator;
 KM dermatological; analgesic; virucide; antibacterial; fungicide.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:37:44 ; Search time 64 Seconds
(without alignments)
969.720 Million cell updates/sec

Title: US-09-900-448-2
Perfect score: 2185
Sequence: 1 MARVLGAPVALGLMSLCWEL.....NAKALPQPNTSLIGCTH 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2139.5	97.9	462	22	AA023933 Human EST encoded
2	456	20.9	83	21	AA000304 Human secreted pro
3	258	11.8	87	23	ABP31577 Human glycoprotein
4	204	9.3	587	21	AAV90502 Murine transmembra
5	204	9.3	587	21	AAV88405 Mouse MT4-MMP (2)
6	202.5	9.3	470	22	AAE10419 Human matrix meta
7	202.5	9.3	470	22	AAE84614 Amino acid sequenc
8	202.5	9.3	470	22	AAE74595 Human macrophage m
9	202.5	9.3	470	22	AAE49982 Human macrophage m

10	202.5	9.3	470	23	AAU91061 Human head and nec
11	202.5	9.3	470	24	ABP96800 Human COPD related
12	202.5	9.3	470	24	ABR48203 Human bladder can
13	202.5	9.3	470	24	ABU56421 Lung cancer-asso
14	202.5	9.3	470	24	ABU56663 Lung cancer-asso
15	202.5	9.3	473	21	AAE43772 Human cancer asoc
16	199	9.1	519	22	AAE10426 Human matrix meta
17	199	9.1	519	22	AAE84619 Amino acid sequenc
18	199	9.1	605	21	AAV90503 Human transmembran
19	199	9.1	605	21	AAO19713 MT4-MMP catalytic
20	199	9.1	606	21	AAV88404 Human MT4-MMP (2)
21	196	9.0	464	19	AAW52135 Rabbit matrix meta
22	195.5	8.9	477	23	AAU91062 Human head and nec
23	192.5	8.8	476	24	ABU03460 Angiogenesis-asso
24	189.5	8.7	475	23	ABU05574 Breast cancer-asso
25	189.5	8.7	475	24	ABU56595 Lung cancer-asso
26	189.5	8.7	476	22	AAE10421 Lung cancer-asso
27	189.5	8.7	476	22	AAE84612 Amino acid sequenc
28	189.5	8.7	476	24	ABR48147 Human bladder can
29	189.5	8.7	476	24	ABP54455 Matrix metalloprot
30	183.5	8.4	464	22	AAE10416 Human MMP-8alt pol
31	183.5	8.4	467	22	AAE10416 Human matrix meta
32	183.5	8.4	467	22	AAE84610 Amino acid sequenc
33	183.5	8.4	467	22	AAE65358 Human neutrophil c
34	181.5	8.3	469	22	AAE10415 Human matrix meta
35	181.5	8.3	469	22	AAE84606 Amino acid sequenc
36	181.5	8.3	469	24	ABR48148 Human bladder can
37	181.5	8.3	469	24	ABU56596 Lung cancer-asso
38	181.5	8.3	469	24	ABU56597 Lung cancer-asso
39	181.5	8.3	469	24	ABU07445 Protein differenti
40	181.5	8.3	469	24	ABP54454 Matrix metalloprot
41	181.5	8.3	492	13	AAE24863 Sequence of pre-pr
42	181.5	8.3	496	22	AAE75509 Human colon cancer
43	180.5	8.3	457	10	AAE93628 Sequence of human
44	180.5	8.3	459	8	AAE70611 Sequence encoded b
45	179.5	8.2	479	23	ABP41861 Human ovarian anti

ALIGNMENTS

RESULT 1	AA023933	standard; Protein; 462 AA.
ID	AA023933	
XX	AA023933;	
AC	12-OCT-2001	(first entry)
XX		
DE	Human EST encoded protein SRQ ID NO: 1458.	
XX		
KW	Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;	
KW	tomato; monkey; dog; sea urchin; expressed sequence tag; EST;	
KW	diagnostics; forensic test; gene mapping; genetic disorder;	
KW	biodiversity; gene therapy; nutrition.	
OS	Homo sapiens.	
XX		
PN	WO200154477-A2.	
XX		
PD	02-AUG-2001.	
XX		
PF	25-JAN-2001; 2001WO-US02687.	
XX		
PR	25-JAN-2000; 2000US-0491404.	
PR	17-JUL-2000; 2000US-0617746.	
PR	03-AUG-2000; 2000US-0631451.	
PR	15-SEP-2000; 2000US-0663870.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YF, Liu C, Zhou F, Qian XB, Wang Z, Chen R, Asundi V;	
PI	Cao Y, Drmanac RA, Zhang J, Werhman T;	

```
Db      548 GCTATACCCAGTAAGGANTTATCCGAAGGGCTGCAGAAAGTAGAGACCCCTCATG 489
Qy      862 GGAATTATCTGGAAGCTGTGGAATGCGGCTTTATCTGCGCTTCTGCGCTCATG 921
Db      488 GGATTTATCCCTGTACTGTGGAATGCGGCTTTATCTGCGCTTCTGCGCTCATG 430
Qy      922 TCATGGCAGAGACGGCGGCTGTGAGTGTGAGACCTGAAGTCAAGAGCCCAAGCCACGNGA 981
Db      429 TCATGGCAGAGACGGCGGCTGTGAGTGTGAGACCTGAAGTCAAGAGCCCAAGCCACGNGA 370
Qy      982 CAGAGCTTCCCTTGGCCCCCATGAGAGGTAGACGAGGCTTGTGATGAAAAAGTCCCTTG 1041
Db      369 CAGAGCTTCCCTTGGCCCCCATGAGAGGTAGACGAGGCTTGTGATGAAAAAGTCCCTTG 310
Qy      1042 GCCCTTAATCTATGTTCCGCCAATGATGATCCGCGCTTGTACTCATCATCATGATGATCCCAATTGT 1101
Db      309 GCCCTTAATCTATGTTCCGCCAATGATGATCCGCGCTTGTACTCATCATCATGATGATCCCAATTGT 250
Qy      1102 ACTGCTACAGTGAATGTGAGAACTGAATGCAAGCCAGGCGCTTCCGCCAATCCCAAGATG 1161
Db      249 ACTGCTACAGTGAATGTGAGAACTGAATGCAAGCCAGGCGCTTCCGCCAATCCCAAGATG 190
Qy      1162 TGAACAAGTCTCTGAGGCTGCACTCATGAGGAGGCTTCTGACATGAGTCTGAGGAGGCTT 1221
Db      189 TGAACAAGTCTCTGAGGCTGCACTCATGAGGAGGCTTCTGACATGAGTCTGAGGAGGCTT 130
Qy      1222 CACTCTCTAATTTCTCTAATAATAAGACAGATGCTTCTTGGCTTCTGACATGAGGAGGCTT 1281
Db      129 CACTCTCTAATTTCTCTAATAATAAGACAGATGCTTCTTGGCTTCTGACATGAGGAGGCTT 70
Qy      1282 CTGACATGAGTCTGAGGCTGAGCCCACTCCCAAGTTTCTCATATAATAAGACAGATGCTT 1341
Db      69 CTGACATGAGTCTGAGGCTGAGCCCACTCCCAAGTTTCTCATATAATAAGACAGATGCTT 10
Qy      1342 CTTCACTTG 1350
Db      9 CTTCACTTG 1
```

Search completed: December 16, 2003, 05:33:59
Job time : 3998.17 secs

RESULT	14
LOCUS	AV653336 631 bp mRNA linear EST 15-JAN-2002
AV653336	GLC Homo sapiens cDNA clone GLCDUH02 3', mRNA sequence.
DEFINITION	AV653336
ACCESSION	AV653336.1 GI:9874350
VERSION	EST.
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 631)
REFERENCE	Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,Y., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z., and Han,Z.
TITLE	Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
JOURNAL	Proc.Natl.Acad.Sci.U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE	21625106
PUBMED	11752456
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel.: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai. Location/Qualifiers
FEATURES	source 1..631 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="GLCDUH02" /tissue_type="corresponding non cancerous liver tissue" /dev_stage="Adult" /lab_host="SOLR" /clone_idb="GLC" /note="Vector: pbluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	138 a 180 c 161 g 149 t 3 others
ORIGIN	
Query Match	16.7%; Score 530.6; DB 9; Length 631;
Best Local Similarity	98.7%; Pred. No.6.5e-148;
Matches	533; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Dy	823 ATCCGAAAGCGCTGGAGAAGAGTGGGAACCCCTCATGGATTATCCTTGACTCTGTGG 882
Db	1 ATCCGAAAGCGCTGGAGAAGAGTGGGAACCCCTCATGGATTATCCTTGACTCTGTGG 60
Dy	883 ATGGGGCTTTATCTGCCCTGGGTCTTCGGCTTCATATCATGCGAAGCGGGCTGT 942
Db	61 ATGGGGCTTTATCTGCCCTGGGTCTTCGGCTTCATATCATGCGAAGCGGGCTGT 120
Dy	943 GTGGAGTGAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1002
Db	121 GTGGAGTGAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Dy	1003 AGAAGTAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1062
Db	181 AGAAGTAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Dy	1063 ATGGTCCCGGCTTGAAGTCAATCATATGATGATGATGATGATGATGATGATGATGATGATG 1122
Db	241 ATGGTCCCGGCTTGAAGTCAATCATATGATGATGATGATGATGATGATGATGATGATGATG 300
Dy	1123 AACTGAATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1182
Db	301 AACTGAATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY	1183	CTCAGTCAGGGGACCTTCTGACATAGAGTCGGCCCTGGACCTTCCTGCTTCATAT	1244
Db	361	CTCAGTCAGGGGACCTTCTGACATAGAGTCGGCCCTGGACCTTCCTGCTTCATAT	420
QY	1243	AAAGCAGATTCCTTCCTTCCTTCCTGACCTGAGGGGCTTCCTGACATAGTCGGCCCTGC	1302
Db	421	AAAGCAGATTCCTTCCTTCCTTCCTGACCTGAGGGGCTTCCTGACATAGTCGGCCCTGC	480
QY	1303	CCGACCTCCGAGTTTCTCATATTAAGACGATTCCTTCCTTCCTGATCAACGAGACC	1365
Db	481	CCGACCTCCGAGTTTCTCATATTAAGACGATTCCTTCCTTCCTGATCAACGAGACC	540
RESULT 15			
AA705094/c			
LOCUS	AA705094	548 bp	mrna
DEFINITION	zj3h10.61 Soares fetal liver spleen INFLS S1 Homo sapiens CDNA		linear
VERSION	clone IMAGE:462499 3'		EST 24-DEC-1997
KEYWORDS	similar to gb:j03048 HEMOPLEXIN PRECURSOR		
ACCESSION	AA705094		
VERSION	AA705094.1		
KEYWORDS	GI:2715012		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 548)		
COMMENT	Hillier, L., Allan, M., Bowles, L., Dubuque, T., Geisel, G., Joet, S.,		
COMMENT	Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin		
COMMENT	, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,		
COMMENT	White, Y., Wyllie, T., Waterston, R. and Wilson, R.		
COMMENT	WashU-NCI human EST Project		
COMMENT	Unpublished		
COMMENT	Contact: Wilson RK		
COMMENT	Washington University School of Medicine		
COMMENT	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
COMMENT	Tel: 314 286 1800		
COMMENT	Fax: 314 286 1810		
COMMENT	Email: est@watson.wustl.edu		
COMMENT	This clone is available royalty-free through LMLT; contact the		
COMMENT	IMAGS Consortium (info@image.llnl.gov) for further information.		
COMMENT	Seq primer: -40m3 fwd. RT from Amersham		
COMMENT	High quality sequence stop: 462.		
FEATURES	Location/Qualifiers		
SOURCE	1..548		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="GDB:3753522"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:462499"		
	/sex="male"		
	/dev_stage="20 week-post conception fetus"		
	/lab_host="DH10B (ampicillin resistant)"		
	/clone_1lb="Soares fetal liver spleen INFLS S1"		
	/note="Organ: Liver and Spleen Vector: pT73D (Pharmacia)		
	with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;		
	This is a subcloned version of the original Soares fetal		
	liver spleen INFLS library. 1st strand cDNA was primed		
	with a Pac I - oligo (dT) primer 15'		
	AACGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 31',		
	double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Pac I and cloned into the Pac I		
	and Eco RI sites of the modified pT73 vector. Library		
	went through one round of normalization. Library		
	constructed by Bento Soares and M.Facima Bonaldo."		
BASE COUNT	137 a 137 c 160 g 113 t		1 others
ORIGIN			
Query Match	16.6%;	Score 529.6;	DB 9; Length 548;
Best Local Similarity	98.9%;	Pred. No. 1.2e-147;	
Matches	543; Conservative	0; Mismatches	5; Indels
			1; Gaps
			1;
QY	802	GCATACCTTACCTGATGAGCGGTTATCCGAAGCGGCTGAGAGAAAGAGTGGGACCCCTCATG	861

Db 905 CTTCACTGAGGACCCACTACTGCGCTGTGACACACCGGAGTGGCTGCATAGCTGCGC 964

QY 689 CATTGCTCATCAGTGGCCCGAGGGCTCTTCCAGCAGTGAAGTCCCTTTCTCGGGAAGA 748

Db 965 CATTGCTCATCAGTGGCCCGAGGGCTCTTCCAGCAGTGAAGTCCCTTTCTCGGGAAGA- 1023

QY 749 AAAAAGCTTATCTGTGCTCAGGCGCACC 773

Db 1024 AAAAAGCTTATCTGTGCTCAGGCGCACC 1048

RESULT 11
LOCUS A1798878 562 bp mRNA linear EST 18-DEC-1999
DEFINITION we93g08.x1 Soares NFL T GBC_S1 Homo sapiens cDNA clone
IMAGE:2348702.3 similar to gb:U03048 HEMOPEXIN PRECURSOR (HUMAN);,
mRNA sequence.

ACCESSION A1798878
VERSION A1798878
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1426 Std Error: 0.00
Seq primer: -400P from Gdbco
High quality sequence stop: 465.
Location/Qualifiers
1. 562
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2348702"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC_S1"
/note="Organ: pooled; Vector: pRTT3P-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung Nhlh19W, testis NHT, and B-cell
NCI CGAP GCB1) were mixed, and 88 circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 134 a 148 c 161 g 119 t

ORIGIN

Query Match 16.8%; Score 536.8; DB 9; Length 562;
Best Local Similarity 99.3%; Pred. No. 7.9e-150;
Matches 560; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 787 TCTTGACAAAGGAGGCTATACCTTGTGTAAGCGGTTATCCGAAGCGGCTGGAAGAAG 846

Db 562 TCTTGACAAAGGAGGCTATACCTTGTGTAAGCGGTTATCCGAAGCGGCTGGAAGAAG 504

QY 847 TCGGAGCCCTCATGAGTATCTGGAAGTGTGAGATGCGGCTTATCTGCGCTGGGT 906

Db 503 TCGGAG-CCCTCATGAGTATCTGGAAGTGTGAGATGCGGCTTATCTGCGCTGGGT 445

QY 907 CTCTCGGCTCATATCATGAGCAGGAGCGGCGTGTGAGCTGGAAGTCAAGAG 966

Db 444 CTCTCGGCTCATATCATGAGCAGGAGCGGCTGTGAGCTGGAAGTCAAGAG 385

QY 967 CCCAAGCCAGCGAGCAGAGCTTCTTGGCCCATGAGAGGTAAGCGAGCTTGTGA 1026

Db 384 CCCAAGCCAGCGAGCAGAGCTTCTTGGCCCATGAGAGGTAAGCGAGCTTGTGA 325

QY 1027 TGGAAAAGTCCCTTGGCCCTTAATGATGTTCCGCAATGGTCCGGCTTGTATCTATCC 1086

Db 324 TGGAAAAGTCCCTTGGCCCTTAATGATGTTCCGCAATGGTCCGGCTTGTATCTATCC 265

QY 1087 ATGTCTCCATTTGTATCTGCTACATGATGTGAGAAATGTGAATGACCAAGGCTTTC 1146

Db 264 ATGTCTCCATTTGTATCTGCTACATGATGTGAGAAATGTGAATGACCAAGGCTTTC 205

QY 1147 CGCAACCCAGAAATGACACATCTCTGGGCTGCACTCACTGAGGGGCTTGTGACATG 1206

Db 204 CGCAACCCAGAAATGACACATCTCTGGGCTGCACTCACTGAGGGGCTTGTGACATG 145

QY 1207 AGCTTGCGCTGGGCCCACTCTGATTCCTCATTAATAAGACAAATGCTTCTGCTTC 1266

Db 144 AGCTTGCGCTGGGCCCACTCTGATTCCTCATTAATAAGACAAATGCTTCTGCTTC 85

QY 1267 TCACTGAGGGGCTTGTGACATGATGTGCGCTGGGCCCACTCTCCAGTTTCTCATTAAT 1326

Db 84 TCACTGAGGGGCTTGTGACATGATGTGCGCTGGGCCCACTCTCCAGTTTCTCATTAAT 25

QY 1327 AAAAGCAGATGCTTCTTCACTTG 1350

Db 24 AAAAGCAGAAAGCTTCTTCACTTG 1

RESULT 12
LOCUS AV661170 602 bp mRNA linear EST 16-JAN-2002
DEFINITION AV661170 GLC Homo sapiens cDNA clone GLCSPD10.3, mRNA sequence.
ACCESSION AV661170
VERSION AV661170.1 GI:9882184
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 602)
AUTHORS Xu X., Huang J., Liu Z., Qian B., Zhu Z., Yan Q., Cai T., Zhang X.,
Xiao H., Qu J., Liu F., Huang Q., Cheng Z., Li N., Du J., Hu W.,
Shen K., Lu G., Fu G., Zhong W., Xu S., Gu W., Huang W., Zhao X.,
Hu G., Gu J., Chen Z., and Han Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

JOURNAL MEDLINE
PUBMED 11752456
CONTACT: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. 602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCSPD10"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

```

QY 837 GAGAGGAATCGGAGCCCTCATGAGTATCTGAGCTGTGAGTGGGCTTATC 896
Db 489 GAAAGGAATCGGAGCCCTCATGAGTATCTGAGCTGTGAGTGGGCTTATC 430
QY 897 TGGCTGTGAGTGTCTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 956
Db 429 TGGCTGTGAGTGTCTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 370
QY 957 AAGCAGGAGCCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1016
Db 369 AAGCAGGAGCCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 310
QY 1017 GCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1076
Db 309 GCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 250
QY 1077 TACCTCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1136
Db 249 TACCTCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 190
QY 1137 AAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1196
Db 189 AAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 130
QY 1197 TTTGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1256
Db 129 TTTGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 70
QY 1257 TCT-TGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1315
Db 69 TCT-TGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 10
QY 1316 TTTCTCTA 1324
Db 9 TTTCTAATA 1

```

```

RESULT 10
AL532235 1051 bp mRNA linear EST 23-MAY-2003
LOCUS AL532235 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM004YJ20 5-PRIME, mRNA sequence.
ACCESSION AL532235
VERSION AL532235.2 GI:3107067
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1051)
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12795728.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 101 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5958.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM004YJ20&cluster=5958.r. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DM004YJ20P1.
Location/Qualifiers
1.1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004YJ20"

```

```

FEATURES
source

```

```

/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 234 a 281 c 303 g 233 t
ORIGIN

```

```

Query Match 16.9%; Score 537; DB 9; Length 1051;
Best Local Similarity 78.3%; Pred. No. 1.2e-149;
Matches 771; Conservative 0; Mismatches 0; Indels 214; Gaps 2;

```

```

QY 2 TCTGAGCTCAGATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 61
Db 65 TCTGAGCTCAGATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 124
QY 62 ATGCTGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 121
Db 125 ATGCTGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 184
QY 122 TGAAGCGAGACCAAGCCAGACCGAGACGTGACGTGACGTGACGTGACGT 181
Db 185 TGAAGCGAGACCAAGCCAGACCGAGACGTGACGTGACGTGACGTGACGT 244
QY 182 TGAAGCTACCAAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 241
Db 245 TGAAGCTACCAAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 304
QY 242 GAAAGTCAAAATGGGACCGGAGTTAATCTCAGAGAGTGAAGAATTTCCCA 301
Db 305 GAAAGTCAAAATGGGACCGGAGTTAATCTCAGAGAGTGAAGAATTTCCCA 354
QY 302 TGTGAGTGTGACATTCGATCAAGTGTCAAGTGTCTTCTGATCAAGGGG 351
Db 365 TGTGAGTGTGACATTCGATCAAGTGTCAAGTGTCTTCTGATCAAGGGG 424
QY 362 CTGGGTATACCTCTCTGAAAAGAAAGAAAGAAAGAAAGTGTCTCAAG 421
Db 425 CTGGGTATACCTCTCTGAAAAGAAAGAAAGAAAGTGTCTCTCAAGTAT 484
QY 422 TCTGGAATCCCATCCCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 481
Db 485 TCTGGAATCCCATCCCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 544
QY 482 TGAAGCGTCTCTTCTTCCA----- 502
Db 545 TGAAGCGTCTCTTCTTCCAAGGTGACCGGAGTGTGAGTGTGAGTGTGAG 604
QY 503 ----- 502
Db 605 CATGAAGAGCGTTCTCTGCGCAGCTGTGGAATCTCTCTCTGCGCTGAG 664
QY 503 ----- 502
Db 665 CCGCTACTACTGCTTCCAGGTTAACCAATTCCTGCGCTTGCACCTGTGAC 724
QY 503 -----AGCCA 508
Db 725 GCCTCCAGGTACCGGAGAGTGTCCGAGACTACTCATGCCCTGCGCGAGAG 784
QY 509 TGAACACAGGAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 568
Db 785 TGAACACAGGAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 844
QY 569 CTGTAGCCCAATCTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 628
Db 845 CTGTAGCCCAATCTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 904
QY 629 CTTCAGTGGAGCCCACTACTGCGGTGTGACACACCGGAGTGTGAGTGTGAG 688

```

AUTHORS	Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X., Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Gu, J., Chen, Z., and Han, Z.
TITLE	Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE	21625106
PUBMED	11752456
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chc.sh.cn This clone is available at CHGC in Shanghai.
FEATURES	Location/Qualifiers
SOURCE	1..704 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="GLAAH08" /tissue_type="corresponding non cancerous liver tissue" /dev_stage="Adult" /lab_host="SOAR" /clone_lib="GLA" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	167 a 181 c 197 g 155 t 4 others
ORIGIN	
Query Match	21.0%; Score 668; DB 9; Length 704;
Best Local Similarity	98.0%; Pred. No. 2.5e-189;
Matches 685; Conservative	0; Mismatches 13; Indels 1; Gaps 1.
Qy	659 CACCAAGCGGGATGGCTGCGATAGCTGGCCCATTTGCTCATAGTAGGCCACAGGGCTTTC 718
Db	703 CACAAAGCCGGAATGGCTGGCATAGCTGGCCCATTTGCTCATAGTAGGCCACAGGGCTTTC 644
Qy	719 AGCAGTGGATGCTGCTTTTCTTCTGGGAAAGAAAATCTT-ATCTGATCCAGGACCCACAG 777
Db	643 AGCAGTGGATGCTGCTTTTCTTCTGGAAGAAAATCTTAATCTGGTCCAGGACCCACAG 584
Qy	778 TATATGTTTCTTCTGACAAAGGAGGCTTATCCTTAGTAAGGGTTATCCGAAGCGGCTG 837
Db	583 TATATGTTTCTTCTGACAAAGGAGGCTTATCCTTAGTAAGGGTTATCCGAAGCGGCTG 524
Qy	838 AGAAGGAAGTGGGACCCCTCATGAGGATTTATCTGGAATCTGTGTGATGCGGCTTTATCT 897
Db	523 AGAAGGAAGTGGGACCCCTCATGAGGATTTATCTGGAATCTGTGTGATGCGGCTTTATCT 464
Qy	898 GCCCTGGGTTCTTCTGGGCTCCATATCATGACAGACGGCGGCTGTGTGTGGCTGACCTGA 957
Db	463 GCCCTGGGTTCTTCTGGGCTCCATATCATGACAGACGGCGGCTGTGTGTGGCTGACCTGA 404
Qy	958 AGTACAGAGCCCAAGCCACGTGAGACAGAGCTTCTTGGGCCCATAGAAAGTAGACGAG 1017
Db	403 AGTACAGAGCCCAAGCCACGTGAGACAGAGCTTCTTGGGCCCATAGAAAGTAGACGAG 344
Qy	1018 CCTTGTGATAGAAAAGTCCCTTGGCCCTTAATCATGTTTCCGCCAATGATCCCGGCTTGT 1077
Db	343 CCTTGTGATAGAAAAGTCCCTTGGCCCTTAATCATGTTTCCGCCAATGATCCCGGCTTGT 284
Qy	1078 ACCTCATCATGATGCCCAATTTGTATCTGTACAGTGAATGTGAGAAATGAAATGACGCA 1137
Db	283 ACCTCATCATGATGCCCAATTTGTACTGTACAGTGAATGTGAGAAATGAAATGACGCA 224
Qy	1138 AGGCGCTTCCGCAACCCCAAGATGACACAGTCTCTGGGCTGACCTCACTGAGAGGGGCT 1197
Db	223 AGGCGCTTCCGCAACCCCAAGATGACACAGTCTCTGGGCTGACCTCACTGAGAGGGGCT 164

QY	1198	TCTGACATGATCTGCGCTGGCCGCCACCTCTAGTTCCTCATATTAAGACAGATTGCT	125
Db	163	TCTGACATGATCTGCGCTGGCCGCCACCTCTAGTTCCTCATATTAAGACAGATTGCT	104
QY	1258	CTTGCGCTTCTCACTGAGGGGCGCTTCTGACATAGATCTGGCGCCCACTCCCACTT	131
Db	103	CTTGCGCTTCTCACTGAGGGGCGCTTCTGACATAGATCTGGCGCCCACTCCCACTT	44
QY	1318	TCTCATATTAAGACAGATTGCTTCTTCACTTGAATCAA	1356
Db	43	TCTCATATTAAGACAGATTGCTTCTTCACTTGAATAAA	5
RESULT 9	877 bp	linear	EST 21-APR-2001
LOCUS	BG196922/c		
DEFINITION	RS116153 Atherysys RAGE Library Homo sapiens CDNA, mRNA sequence.		
ACCESSION	BG196922		
VERSION	BG196922.1	GI:13718609	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McEllisot,K., Booser,S., Mays,R., Smith,E., Veloso,N., Kikla,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Darrig,J. and Ducar,M.		
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression		
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)		
MEDLINE	21227151		
PUBMED	11329013		
COMMENT	Contact: Scott J. Cain Atherysys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scain@atherysys.com		
FEATURES	High quality sequence stop: 326. Location/Qualifiers 1..877 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cell_line="HT1080" /clone_id="Atherysys RAGE Library" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."		
BASE COUNT	206 a 213 c 289 g 169 t		
ORIGIN			
Query Match	17.2%; Score 549; DB 10; Length 877;		
Best Local Similarity	90.7%; Pred. No. 2,46-153;		
Matches	607; Conservative 0; Mismatches 60; Indels 2; Gaps 2;		
QY	658	ACACGACCGGGAATGGCTGGCATGATGCGCCATTGCTCATCAGTGGCCCGAGGATCTT	717
Db	669	ACCAACCTTGGGAGGCTTGAATGAGGACACATGTTTCATCAGGGGCCCAAGTCCCT	610
QY	718	CAGCAGTGAATGCTGCTTTTCTTCTGGAGAG-AAAACCTATCTGTTCAGGAGCCAG	776
Db	609	AGCAGTGAAGTCTCCTCTTCTGGAGACATAAACCTTATTTGGTTCAAGGGCCCCAG	550
QY	777	GTAATATGCTTCTCGACAAAGGAGGCTATACCTTAAAGCGGTATCCGAAACGGCTG	836
Db	549	GTATATTTTCTCGCTCAAGGAGGCTATCCCTTATTAACCGGTATCTCAAGCGCTG	490

Db	451	ATGGGATTATTCCTGGACTCTGTGGATGGGGCTTTATCTGGCCCTGGCTTCTCGGCTCC	392
Qy	919	ATATCATGGCAGAGACGGCGGCTGTGTGTGGCTGACCTGAAGTCAAGAGCCCAAGCAGT	978
Db	391	ATATCATGGCAGAGACGGCGGCTGTGTGTGGCTGACCTGAAGTCAAGAGCCCAAGCAGT	332
Qy	979	GGAAGAGGCTTCCTTGGGCCCCATGAGAAAGGTAGACGAGCCTTGTGTATGAGAAAAGTCCC	1038
Db	331	GGAAGAGGCTTCCTTGGGCCCCATGAGAAAGGTAGACGAGCCTTGTGTATGAGAAAAGTCCC	272
Qy	1039	TTGGCCCTTAATCACTAGTTCCGGCAATGGTCCGGCTGTACTCATCAATGATGCCAATT	1098
Db	271	TTGGCCCTTAATCACTAGTTCCGGCAATGGTCCGGCTGTACTCATCAATGATGCCAATT	212
Qy	1099	TGTACTGCTACAGTGAATGTGGAGAAACTGAATGACGCAAGGCCCTTCCGAAACCCGAGA	1158
Db	211	TGTACTGCTACAGTGAATGTGGAGAAACTGAATGACGCAAGGCCCTTCCGAAACCCGAGA	152
Qy	1159	ATGTGACCAAGTCTCCCTGGGCTGCATCACTGAGGGGCTTCTGACATGAGTCTGGCTGG	1218
Db	151	ATGTGACCAAGTCTCCCTGGGCTGCATCACTGAGGGGCTTCTGACATGAGTCTGGCTGG	92
Qy	1219	CCCCAAGCTCTAGTTCCTCATTAATTAAGACAGATTGCTTTCGCTTCTCACTAGGGGGC	1278
Db	91	CCCCAAGCTCTAGTTCCTCATTAATTAAGACAGATTGCTTTCGCTTCTCACTAGGGGGC	32
Qy	1279	CTTCTGACATGAGTCTGGGCTGGGCCCCAC	1307
Db	31	CTTCTGAAATAGATATATATATTAATAAAAAM	3

FEATURES	source
LOCUS	1156 bp mRNA linear EST 22-MAY-2003
DEFINITION	BX448764 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
ACCESSION	BX448764
VERSION	CS0DM011CB23 3-PRIME, mRNA sequence.
KEYWORDS	BX448764.1 GI:31021961
ORGANISM	EST.
	Homo sapiens (human)
	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 1156)
TITLE	L1,W.B., Gruber,C., Jeessee,J. and Polayes,D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished
	Contact: Genoscope
	Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - France
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
	Library was constructed by life technologies, a division of
	Invitrogen. This sequence belongs to sequence cluster 5958.r For
	more information about this cluster, see
	http://www.genoscope.cns.fr/
	cg1-bin/cluster.cgi?seq=CS0AM011CA12NP1&cluster=5958.r. Contact :
	Feng Liang Email : fliang@lifetech.com URL :
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
	Faraday Avenue Genoscope sequence ID : CS0AM011CA12NP1.
	Location/Qualifiers
	1..1156

```

1..1156
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM011YB23"
/tissue_type="FETAL_LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL_LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6

```

```
vector. Library was not normalized."
BASE COUNT      264 a      260 c      301 g      240 t      91 others
ORIGIN
```

Query Match	21.6%	Score 688.8;	DB 13;	Length 1156;
Best Local Similarity	97.3%	Pred. No. 2e-195;		
Matches 745; Conservative	4;	Mismatches 12;	Indels 5;	Gaps 5;

OY	503	AGGCAATGGAACAAGAAATGGAGCTGGCCATGTGGAAACAGTAACCAACATATGGCCCTGAGTA	562
Db	765	AGSCATAGGAACAAGAAATGGAACTGGCCATGGGAACATACCAACATATGGCCCTGAGTA	706
OY	563	TATGCGCTGTAAGCCCACTATAGTCTTGTCTGCACTGACGCTTGACAAACATAGGTGCAC	622
Db	705	TATGCGCTKTAAAGCCCACTATAGTCTTGTCTGCAAC - AAGCTTGACAAACAT - GTGCGAC	648
OY	623	CTATGCTTCAATGGGAACCACTAATGGGCTCTGGACAACAAGCCGGATATGGCTGGCATAG	682
Db	647	CTATGCTTCAAGT - GGAACCACTAATGGGCTCTGGACAACAAGCCGGATATGGCTGGCATAG	589
OY	683	CTGGCCCAATGGCTCATAGTGGCCCCCAAGGATCCCTTCAGAGTGGATCTGCTTTTCTGT	742
Db	588	CTGGCCCAATGGCTCATAGTGGCCCCCAAGG - CCTTCAGAGTGGATCTGCACTTTCTGT	530
OY	743	GGAAAGAAAACTATATGTGTCCAGGGCACCCAGATATATGTCTTCTGCAAAAGGAGG	802
Db	529	GGAAAGAAAACTATATGTGTCCAGGGCACCCAGGTATATGTCTTCTTCAAAAGGAGG	470
OY	803	CTATACCTCTAGTAAGCGGTTATCCGAAGCGGCTGGAAAGAAAGTGGGAAACCCCTCATGG	862
Db	469	CTATACCTCTAGTAAGCGGTTATCCGAAGCGGCTGGAAAGAAAGTGGGAAACCCCTCATGG	410
OY	863	GATTATCCTGGACTCTGTGATGTCGGGCTTTATCTGTGCTCTGGGCTTCCATAT	922
Db	409	GATTATCCTGGACTCTGTGATGTCGGGCTTTATCTGTGCTCTGGGCTTCCATAT	350
OY	923	CATGGCAGAACCGCGGCTGTGGTGGCTGACCTTGAACTCAGAAAGCCCAAGCCAGTGGAC	982
Db	349	CATGGCAGAACCGCGGCTGTGGTGGCTGGAACCTTGAACTCAGAAAGCCCAAGCCAGTGGAC	290
OY	983	AGACTTCTTTGGGCCCATAGAAAGGTAAACGAGCCTTGTGTTATGGAAAAAGTCCCTTGG	1042
Db	289	AGACTTCTTTGGGCCCATAGAAAGGTAAACGAGCCTTGTGTTATGGAAAAAGTCCCTTGG	230
OY	1043	CCCTAATCATGTTCGCGCAATGGTCCCGGCTTGATCCTCATGCATATGTCCTCAATTTGTA	1102
Db	229	CCCTAATCATGTTCGCGCAATGGTCCCGGCTTGATCCTCATGCATATGTCCTCAATTTGTA	170
OY	1103	CTGCTACAGTATGTGAGAACTGAATGCAAGCCAAAGCCCTTCCGCAACCCCAAGAAATG	1162
Db	169	CTGCTACAGTATGTGAGAACTGAATGCAAGCCAAAGCCCTTCCGCAACCCCAAGAAATG	110
OY	1163	GACCAATCTCTGGGGCTGCACTCATGAGGGGCTTTCTGACATGAAATCTGGCTTGGCCCC	1222
Db	109	GACCAATCTCTGGGGCTGMACTCATGAGGGGCTTTCTGACATGAAATCTGGCTTGGCCCC	50
OY	1223	ACGCTCCTAGTTC - CTCATAATTAAGACAGATGCTTCTTGGCTCT 1267	
Db	49	ACGCTCCTAGTTCCTCATTAATTAAGAAATGTTTCTTGGCTTTT 4	

RESULT 8	
AV645403/c	
LOCUS	AV645403 704 bp mRNA 11near EST 15-JAN-2002
DEFINITION	AV645403 GLA Homo sapiens cDNA clone G1AAAH08 3', mRNA sequence.
ACCESSION	AV645403
VERSION	AV645403.1 GI:9866417
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 704)

```

/tissue_type="FETAL BRAIN"
/dev_stage="Fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dt) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

BASE COUNT 243 a 306 c 340 g 268 t 44 others

Query Match 23.7%; Score 756.6; DB 13; Length 1201;
 Best Local Similarity 94.9%; Pred. No. 7.4e-216;
 Matches 787; Conservative 13; Mismatches 27; Indels 2; Gaps 2;

```

QY 503 AGGCGATGACACAGAAATGAGTGGAGTGGGACATGGGAAACAGTACCACCATGGCCCTGAGTA 562
DB 272 AGGCGATGACACAGAAATGAGTGGAGTGGGACATGGGAAACAGTACCACCATGGCCCTGAGTA 331
QY 563 TATGCGCTGAGGCCCATCATGCTTGTCTGCACTGACCTGACAAACATGGTGGCAC 622
DB 332 TATGCGCTGAGGCCCATCATGCTTGTCTGCACTGACCTGACAAACATGGTGGCAC 391
QY 623 CTATGCTTCACTGAGGACCACTACTGCGCTGAGACACGACCGGAGTGGCTGACATAG 682
DB 392 CTATGCTTCACTGAGGACCACTACTGCGCTGAGACACGACCGGAGTGGCTGACATAG 451
QY 683 CTGCGCCATGCTCATCATGAGGCGCCAGGGTCTTTCAGCAGTGGATGCTCCCTTTCTCG 742
DB 452 CTGCGCCATGCTCATCATGAGGCGCCAGGGTCTTTCAGCAGTGGATGCTCCCTTTCTCG 511
QY 743 GGAAGAAAACTCTATGCTGTCAGGAGGACCCAGTATGCTTCTCTGCAAAAGGAGG 802
DB 512 GGAAGAAAACTCTATGCTGTCAGGAGGACCCAGTATGCTTCTCTGCAAAAGGAGG 571
QY 803 CTATACCTCTAGTAAAGCGGTTATCCGAAAGCGCTGAGAAAGAAAGTCCGACCTCATAG 862
DB 572 CTATACCTCTAGTAAAGCGGTTATCCGAAAGCGCTGAGAAAGAAAGTCCGACCTCATAG 631
QY 863 GATTATCTGAGACTCTGTGAGATGGCGCTTATATCTGCGCTGCTTCTGCTGCTCATAT 922
DB 632 GATTATCTGAGACTCTGTGAGATGGCGCTTATATCTGCGCTGCTTCTGCTGCTCATAT 691
QY 923 CATGAGACGAGCGGCTGTGAGTGGCTGAGACCTGAAATCGAGAGCCCAAGCCACGTGGAC 982
DB 692 CATGAGACGAGCGGCTGTGAGTGGCTGAGACCTGAAATCGAGAGCCCAAGCCACGTGGAC 751
QY 983 AGAGCTTCTTGGGCCCATGAGAAAGTGAACGAGACCTTGTGATGAGAAAGTCCCTTGG 1042
DB 752 AGAGCTTCTTGGGCCCATGAGAAAGTGAACGAGACCTTGTGATGAGAAAGTCCCTTGG 811
QY 1043 CCTTAATCATGTTCCGCAATGAGTCCGCGCTGTATCTCATTCATGCTGCCAATTTGTA 1102
DB 812 CCTTAATCATGTTCCGCAATGAGTCCGCGCTGTATCTCATTCATGCTGCCAATTTGTA 871
QY 1103 CTGCTACAGTATGAGTGAAGAACTGATGAGCCCAAGCCCTTCCGAAACCCCAAGATGT 1162
DB 872 CTGCTACAGTATGAGTGAAGAACTGATGAGCCCAAGCCCTTCCGAAACCCCAAGATGT 931
QY 1163 GACCAAGTCTCTGGGCTGACATCACTG-AGGGGCTTCTGACATGATGCTGGCCCTG 1221
DB 932 GACCAAGTCTCTGGGCTGACATCACTG-AGGGGCTTCTGACATGATGCTGGCCCTG 991
QY 1222 CACC-TCCTAGTTCTCATATAAAGACAGATTGCTTCTGCTTCTCACTAGAGGGCT 1280
DB 992 CACC-TCCTAGTTCTCATATAAAGACAGATTGCTTCTGCTTCTCACTAGAGGGCT 1051
QY 1281 TCTGACATGATGCTGGGCTGGCCCACTCTCCAGTTTCTCATATAA 1329
DB 1052 TCTTGAATAGTGTGSCCTGGCCCACTCTCCGAGTTTATATAAARMA 1100

```

RESULT 6

```

BX405876/c
LOCUS 1201 bp mRNA linear EST 13-MAY-2003
DEFINITION BX405876 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
ACCESSION CS0DM009Y021.3-PRIME, mRNA sequence.
VERSION BX405876
KEYWORDS BX405876.1 GI:30652298
SOURCE EST.
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5958.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AM009AH1NP1&cluster=5958.r. Contact :
Feng Liang Email: fliang@life.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CS0AM009AH1NP1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM009Y021"
/tissue_type="FETAL LIVER"
/dev_stage="Fetal"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dt) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

BASE COUNT 245 a 303 g 265 t 120 others

```

QY 503 AGGCGATGACACAGAAATGAGTGGAGTGGGACATGGGAAACAGTACCACCATGGCCCTGAGTA 562
DB 811 AGGCGATGACACAGAAATGAGTGGAGTGGGACATGGGAAACAGTACCACCATGGCCCTGAGTA 752
QY 563 TATGCGCTGAGGCCCATCATGCTTGTCTGCACTGACCTGACAAACATGGTGGCAC 622
DB 752 TATGCGCTGAGGCCCATCATGCTTGTCTGCACTGACCTGACAAACATGGTGGCAC 692
QY 623 CTATGCTTCACTGAGGACCACTACTGCGCTGAGACACGACCGGAGTGGCTGACATAG 682
DB 692 CTATGCTTCACTGAGGACCACTACTGCGCTGAGACACGACCGGAGTGGCTGACATAG 632
QY 683 CTGCGCCATGCTCATCATGAGGCGCCAGGGTCTTTCAGCAGTGGATGCTCCCTTTCTCG 742
DB 632 CTGCGCCATGCTCATCATGAGGCGCCAGGGTCTTTCAGCAGTGGATGCTCCCTTTCTCG 572
QY 743 GGAAGAAAACTCTATGCTGTCAGGAGGACCCAGTATGCTTCTCTGCAAAAGG 799
DB 572 GGAAGAAAACTCTATGCTGTCAGGAGGACCCAGGATGATGCTTCTCTGCAAAAGG 512
QY 800 -AGGCTATACCTTATGAGGCTTATCCGAAAGGCTGAGAAAGAAAGTCCGACCTTC 858
DB 511 TAGGCTATACCTTATGAGGCTTATCCGAAAGGCTGAGAAAGAAAGTCCGACCTTC 452
QY 859 ATGGATTAATCCGAGACTGTGAGATGGCGCTTATATGCTGCGGCTTC 918

```

QY 1283 TGACATGAGTCTGCGCCGACCCACCTCC-CCAGTTTCATATAAAGACAGATTGCT 1340
 Db 66 TGACATGAGTCTGCGCCGACCCACCTCCAGTTTCATATAAAGANNANNNMT 8

RESULT 4
 AL564746/c
 LOCUS AL564746 Homo sapiens PBTAL LIVER Homo sapiens cDNA clone EST 31-MAY-2003
 DEFINITION AL564746 Homo sapiens PBTAL LIVER Homo sapiens cDNA clone
 CS0DM007YC08 3-PRIME, mRNA sequence.
 ACCESSION AL564746
 VERSION AL564746.2 GI:31288721
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 Li W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 16, 2001 this sequence version replaced GI:12915461.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5958.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DM007B04NP1&cluster=5958.r. Contact :
 Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DM007B04NP1.
 Location/Qualifiers
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DM007YC08"
 /isue_type="PBTAL LIVER"
 /dev_stage="fetal"
 /note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."
 vector. Library was not normalized."
 BASE COUNT 278 a 291 c 323 g 267 t 42 others
 ORIGIN

Query Match 25.2%; Score 802.6; DB 9; Length 1201;
 Best Local Similarity 98.8%; Pred. No. 9.8e-230;
 Matches 825; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

QY 503 AGCCATGACACAGAGATGGAAGTGGCCATGGGAAAGTACCCACCATGGCCCTGAGTA 562
 Db 836 AGCCATGACACAGAGATGGAAGTGGCCATGGGAAAGTACCCACCATGGCCCTGAGTA 777

QY 563 TATGCGCTGAGCCACATCTAGCTTGTGTGACGTGAGCTGTGACAAACATGAGTGCAC 652
 Db 776 TATGCGCTGAGCCACATCTAGCTTGTGTGACGTGAGCTGTGACAAACATGAGTGCAC 717

QY 623 CTATGCTTCACTGAGGACCCACTAGCGCTGTGACACACGCGGATGGCTGAGTAG 682
 Db 716 CTATGCTTCACTGAGGACCCACTAGCGCTGTGACACACGCGGATGGCTGAGTAG 657

QY 683 CTGCGCCATGCTCATGAGTGGCCCGGCTCTTCAAGCATGAGTGTCTCTTCTCTG 742
 Db 656 CTGCGCCATGCTCATGAGTGGCCCGGCTCTTCAAGCATGAGTGTCTCTTCTCTG 597

QY 743 GGAAGAAAACTCTATCTGCTCAAGGACCCAGGTATATGCTTCTCTGCAAAAGGAGG 802

Db 596 GGAAGAAAACTCTATCTGCTCAAGGACCCAGGTATATGCTTCTCTGCAAAAGGAGG 537
 QY 803 CTATACCTTATGTAAGCGGTTATCCGAGGCGCTGAGAAAGATGCGGACCCCTCATG 862
 Db 536 CTATACCTTATGTAAGCGGTTATCCGAGGCGCTGAGAAAGATGCGGACCCCTCATG 477

QY 863 GATTATCCGGAAGCTGATGAGTGGCCCTTATCTGCGCTGAGTCTTCTGCGCTCATAT 922
 Db 476 GATTATCCGGAAGCTGATGAGTGGCCCTTATCTGCGCTGAGTCTTCTGCGCTCATAT 417

QY 923 CATGCGAGACGCGCGCTGTGTGAGCTGAGACCTGAGTGAAGGCCCAAGCCAGTGAAC 982
 Db 416 CATGCGAGACGCGCGCTGTGTGAGCTGAGACCTGAGTGAAGGCCCAAGCCAGTGAAC 357

QY 983 AAGACTTCTTGGCCCATGAGAGGTGAGCGAGCTTGTATGAGAAAGTCCCTTGG 1042
 Db 356 AAGACTTCTTGGCCCATGAGAGGTGAGCGAGCTTGTATGAGAAAGTCCCTTGG 297

QY 1043 CCTTAATGATGTCGCGCAATGATGTCGCGCTTGTACCTCATGCAATGTCCTCAATTGTA 1102
 Db 296 CCTTAATGATGTCGCGCAATGATGTCGCGCTTGTACCTCATGATGTCCTCAATTGTA 237

QY 1103 CTGCTACAGTATGTAAGAACTGATGACGACCAAGCCCTTCCGCAACCCAGATGT 1162
 Db 236 CTGCTACAGTATGTAAGAACTGATGACGACCAAGCCCTTCCGCAACCCAGATGT 177

QY 1163 GACCACTTCTGCGCTGCTGACCTGACCTGAGGCGCTTGTACATGAGTCTGCGCTGCCCC 1222
 Db 176 GACCACTTCTGCGCTGCTGACCTGACCTGAGGCGCTTGTACATGAGTCTGCGCTGCCCC 117

QY 1223 ACCTCTATGTCCTATATAAAGACAGATTGCTTGTGCTTCTGACGAGGCGCTTC 1282
 Db 116 ACCTCTATGTCCTATATAAAGACAGATTGCTTGTGCTTCTGACGAGGCGCTTC 57

QY 1283 TGACATGAGTCTGCGCCGACCCACCTCCAGTTTCATATAAAGACAGATTGCT 1337
 Db 56 TGACATGAGTCTGCGCCGACCCACCTCCAGTTTCATATAAANNANNNMT 4

RESULT 5
 BX440692
 LOCUS BX440692 Homo sapiens PBTAL BRAIN Homo sapiens cDNA clone
 DEFINITION CS0DF010YK01 5-PRIME, mRNA sequence.
 ACCESSION BX440692
 VERSION BX440692.1 GI:30781865
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 Li W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5958.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DF010AF010P1&cluster=5958.r. Contact :
 Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DF010AF010P1.
 Location/Qualifiers
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF010YK01"

```

Db      607 TTTCCTGGGAGAGAACTATCTATGTCAGGAGCCAGGATATGTCTTCTGACAA 548
Qy      796 AGGAGAGCTATACCTAGTAAAGGCTTATCCGAAGCGCTGGAGAGAGTGGGACCC 855
Db      547 AGGAGAGCTATACCTAGTAAAGGCTTATCCGAAGCGCTGGAGAGAGTGGGACCC 488
Qy      856 CTGATGAGATATCTGAGACTCTGTGATGAGCGGCTTATCTGCGCTTCTTCTGAG 915
Db      487 CTGATGAGATATCTGAGACTCTGTGATGAGCGGCTTATCTGCGCTTCTTCTGAG 428
Qy      916 TCATATCATGAGAGAGCGGCGCTGTGTGCTGAGACTGAAATCGAGAGCCCAAGCCA 975
Db      427 TCATATCATGAGAGAGCGGCGCTGTGTGCTGAGACTGAAATCGAGAGCCCAAGCCA 368
Qy      976 CGTGAAGAGCTTCTTGGCCCCCAGAGAGAGTAAAGAGCGCTGTGTATGAGAAAGT 1035
Db      367 CGTGAAGAGCTTCTTGGCCCCCAGAGAGAGTAAAGAGCGCTGTGTATGAGAAAGT 308
Qy      1036 CCCTTGGCCCTTACTCATGTTTCCGCCAATGTCGCCGCTTGTACCTCATGATGTCCTCA 1095
Db      307 CCCTTGGCCCTTACTCATGTTTCCGCCAATGTCGCCGCTTGTACCTCATGATGTCCTCA 248
Qy      1096 ATTGTACTGCTACAGTATGTGAGAGAACTGAAATGAGCCAAAGCCCTTCCGCAACCC 1155
Db      247 ATTGTACTGCTACAGTATGTGAGAGAACTGAAATGAGCCAAAGCCCTTCCGCAACCC 188
Qy      1156 AGAATGTGACCAATCTCTGGGCTGCACTACAGAGGGGCTTGTACATGATGTGGCC 1215
Db      187 AGAATGTGACCAATCTCTGGGCTGCACTACAGAGGGGCTTGTACATGATGTGGCC 128
Qy      1216 TGGCCCCCAGCTCTAGTCTCTCAATTAATAAGACAGATTGCTTCTGCTTCTGATGAG 1275
Db      127 TGGCCCCCAGCTCTAGTCTCTCAATTAATAAGACAGATTGCTTCTGCTTCTGATGAG 68
Qy      1276 GGGCTTCTGACATGAGTCTGGGCTGGCCCACTCCCACTTCTCATTAATAAGACAG 1335
Db      67 GGGCTTCTGACATGAGTCTGGGCTGGCCCACTCCCACTTCTCATTAATAAGACAG 9
Qy      1336 TTGCTTCT 1343
Db      8 TTGCTTCT 1

```

```

RESULT 3
AL564329/c 1053 bp mRNA linear EST 31-MAY-2003
LOCUS AL564329 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM004YJ20 3-PRIME, mRNA sequence.
ACCESSION AL564329
VERSION AL564329.2 GI:31288307
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Li,W.B., Gruber,C., Jessee,J. and Polyes,D.
Full-length cDNA libraries and normalization.
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12914625.
Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5958.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM004DE10NP1&cluster=5958.r. Contact :
Feng Liang Email: fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DM004DE10NP1.

```

```

FEATURES
source
Location/Qualifiers
1..1053
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004YJ20"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_1ib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT.6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 237 a 286 c 306 g 210 t 14 others
ORIGIN
Query Match 25.6%; Score 816; DB 9; Length 1053;
Best Local Similarity 98.6%; Pred. No. 8e-234;
Matches 827; Conservative 3; Mismatches 8; Indels 1; Gaps 1;
Qy 503 AGGCAATGACACAGAAATGGGACTGGCCATGGGAAACAGTACCACATGGCCCTGAGTA 562
Db 846 AGGCCATGACACAGAAATGGGACTGGCCATGGGAAACAGTACCACATGGCCCTGAGTA 787
Qy 563 TATGGCTGAGCCCATCTAGTCTGTCTGACTGAGAGCTGACACCATGGTGGCCAC 622
Db 786 TATGGCTGAGCCCATCTAGTCTGTCTGACTGAGAGCTGACACCATGGTGGCCAC 727
Qy 623 CTATGCTTCACTGAGGACCACTACTGAGGCTGACACACAGCCGGAGTGGCTGAGTA 682
Db 726 CTATGCTTCACTGAGGACCACTACTGAGGCTGACACACAGCCGGAGTGGCTGAGTA 667
Qy 683 CTGAGCCATGCTCATCATGAGGCCCCAGAGGCTCTTGAAGAGTGAATGCTGCTTCTG 742
Db 666 CTGAGCCATGCTCATCATGAGGCCCCAGAGGCTCTTGAAGAGTGAATGCTGCTTCTG 607
Qy 743 GGAAGAAAACCTATCTGCTCAAGGACCCAGGATATGCTTCTGACAAAGGAGG 802
Db 606 GGAAGAAAACCTATCTGCTCAAGGACCCAGGATATGCTTCTGACAAAGGAGG 547
Qy 803 CTATACCTAGTAAAGGCTTATCCGAAGCGGCTGAGAGAGTGGGACCCCTCATG 862
Db 546 CTATACCTAGTAAAGGCTTATCCGAAGCGGCTGAGAGAGTGGGACCCCTCATG 487
Qy 863 GATTATCTGAGACTCTGTGAGTGGGCTTATCTGCTGAGTCTTCTGCTGCTCATAT 922
Db 486 GATTATCTGAGACTCTGTGAGTGGGCTTATCTGCTGAGTCTTCTGCTGCTCATAT 427
Qy 923 CATGAGAGAGCGGCGCTGTGAGCTGAGACCTGAAGTCAAGAGCCCAAGCGAGAC 982
Db 426 CATGAGAGAGCGGCGCTGTGAGCTGAGACCTGAAGTCAAGAGCCCAAGCGAGAC 367
Qy 983 AGAGCTTCTTGGCCCATGAGAGGTAAGCGAGCTGTGTATGAGAAAGTCCCTTGG 1042
Db 366 AGAGCTTCTTGGCCCATGAGAGGTAAGCGAGCTGTGTATGAGAAAGTCCCTTGG 307
Qy 1043 CCTTAACTCATGTTCCGCAATGATGTCGGCTTGAATCTCATCATGATGTCCTCAATTGTA 1102
Db 306 CCTTAACTCATGTTCCGCAATGATGTCGGCTTGAATCTCATCATGATGTCCTCAATTGTA 247
Qy 1103 CTGCTACAGTAAATGAGAACTGAATGACCAAGGCGCTTCCGCAACCCAGAAATG 1162
Db 246 CTGCTACAGTAAATGAGAACTGAATGACCAAGGCGCTTCCGCAACCCAGAAATG 187
Qy 1163 GACCACTCTCTGGGCTGACTCATGAGAGGCTTCTGACATGAGTGTGGCTGGCCCC 1222
Db 186 GACCACTCTCTGGGCTGACTCATGAGAGGCTTCTGACATGAGTGTGGCTGGCCCC 127
Qy 1223 ACTCTTACTTCTCATTAATAAGACAGATTGCTTCTGCTTCTCATGAGGAGGCTTTC 1282
Db 126 ACTCTTACTTCTCATTAATAAGACAGATTGCTTCTGCTTCTCATGAGGAGGCTTTC 67

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 10:53:38 ; Search time 3986.5 Seconds
(without alignments)
19424.068 Million cell updates/sec

Title: US-09-900-448-1

Perfect score: 3186
Sequence: 1 cctcgcagctcagcatgct.....accttttcgtcgttcctca 3186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdg:*
26: em_gss_pbg:*
27: em_gss_vrt1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	834.2	26.2	992	13	BX325177 BX325177
C 2	822.4	25.8	1201	9	AL534854 AL534854
C 3	816	25.6	1053	9	AL564329 AL564329
C 4	802.6	25.2	1201	9	AL564746 AL564746

5	756.6	23.7	1201	13	BX440692	BX440692
C 6	743.4	23.3	1201	13	BX405876	BX405876
C 7	688.8	21.6	1156	13	BX448764	BX448764
C 8	668	21.0	704	9	AV645403	AV645403
C 9	549	17.2	877	10	BG196922	BG196922
C 10	537	16.9	1051	9	AL532235	AL532235
C 11	536.8	16.8	562	9	AI798878	AI798878
C 12	533.2	16.7	602	9	AV661170	AV661170
C 13	531.2	16.7	914	13	BX412777	BX412777
C 14	530.6	16.7	631	9	AV553336	AV553336
C 15	529.6	16.6	548	9	AA705094	AA705094
C 16	527.8	16.6	551	14	CB154152	CB154152
C 17	506.4	15.9	582	9	AI133162	AI133162
C 18	505.2	15.9	542	2	HSN087830	HSN087830
C 19	505.2	15.9	542	14	CB111971	CB111971
C 20	505.2	15.9	546	14	CB120347	CB120347
C 21	504.8	15.8	1201	13	BX448765	BX448765
C 22	504.2	15.8	1061	12	BM921160	BM921160
C 23	503.6	15.8	540	14	CB113371	CB113371
C 24	503.6	15.8	657	14	CB141709	CB141709
C 25	502.2	15.8	989	13	BX325178	BX325178
C 26	500.8	15.7	569	9	AV661177	AV661177
C 27	500.4	15.7	548	14	CB147228	CB147228
C 28	499.8	15.7	1201	13	BX463236	BX463236
C 29	494.2	15.5	693	10	BG617908	BG617908
C 30	484	15.5	630	10	AW965055	AW965055
C 31	491.6	15.4	791	10	BG617459	BG617459
C 32	489.4	15.4	570	10	BG616601	BG616601
C 33	486.6	15.3	657	12	BM780523	BM780523
C 34	486.2	15.3	545	14	TS4902	TS4902
C 35	482.8	15.2	1201	9	AL564442	AL564442
C 36	481.4	15.1	498	14	CB147209	CB147209
C 37	479.8	15.1	488	14	CB154135	CB154135
C 38	478.8	15.0	482	14	CB154307	CB154307
C 39	465.8	14.6	485	9	AA514026	AA514026
C 40	460.6	14.5	1201	13	BX405877	BX405877
C 41	459	14.4	784	14	CB165194	CB165194
C 42	457.6	14.4	503	14	TT1372	TT1372
C 43	456	14.3	749	9	AV693145	AV693145
C 44	456	14.3	890	10	BF126689	BF126689
C 45	454.6	14.3	620	12	BM780657	BM780657

ALIGNMENTS

RESULT 1
LOCUS BX325177/c
DEFINITION BX325177 Homo sapiens T CELLS (JURKAT CELL LINE) COI 10-NORMALIZED
VERSION BX325177
KEYWORDS BX325177.1 GI:30340447
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 992)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5958.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0A0J010BC12NP1c1cluster=5958.r. Contact :
Feng Liang Email : fliang@lifeitech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Db 361 GAGCAAGACCAGACCTTCCCGAGGCCAGAGATGAACAGCTCT 408

RESULT 15

US-09-918-995-7738
; Sequence 7738, Application US/09918995
; Publication No. US2003073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7738
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7738

Query Match 8.7%; Score 278; DB 11; Length 384;
Best Local Similarity 87.7%; Pred. No. 2.2e-81;
Matches 335; Conservative 0; Mismatches 0; Indels 47; Gaps 1;

QY 1 CTCTGCAGCTCAGCATGGCTAGGGTACTGGAGGACCCGTTGCACTGGGGTTGTGAGCC 60
Db 3 CTCTGCAGCTCAGCATGGCTAGGGTACTGGAGGACCCGTTGCACTGGGGTTGTGAGCC 62
QY 61 TATGCTGCTCTGGCCATTGGCCACCCCTCTTCCTCC----- 97
Db 63 TATGCTGCTCTGGCCATTGGCCACCCCTCTTCCTCCGTAAGTAAAGCTGGGACTAGAG 122
QY 98 -----GACTAGTGGCCATGGAGATGTTGCTGAAGGCCGAGAC 133
Db 123 CGAAGATTGAGTTCTGGGCTAGGGACTAGTGGCCATGGAGATGTTGCTGAAGGCCGAGAC 182
QY 134 CAAGCCAGACCCAGACGCTGACTGAACGCTGCTCAGATGCTGAGAGCTTTGATGCTACAC 193
Db 183 CAAGCCAGACCCAGACGCTGACTGAACGCTGCTCAGATGCTGAGAGCTTTGATGCTACAC 242
QY 194 CCTGATGACAATGGAACCAATGCTGTTTAAAGGGAGTTTGTGGAAGAGTCAAA 253
Db 243 CCTGATGACAATGGAACCAATGCTGTTTAAAGGGAGTTTGTGGAAGAGTCAAA 302
QY 254 ATGGGACCGGAGTTAATCTCAGAGAGATGGAAGATTTCCAGCCCTGTGATGCTGC 313
Db 303 ATGGGACCGGAGTTAATCTCAGAGAGATGGAAGATTTCCAGCCCTGTGATGCTGC 362
QY 314 ATTCCGTCAAGGTCAACAAGT 335
Db 363 ATTCCGTCAAGGTCAACAAGT 384

Search completed: December 16, 2003, 06:45:19
Job time : 676.145 secs

Matches 468; Conservative 0; Mismatches 29; Indels 15; Gaps 13;

QY 839 GAAGAAGTGGGACCCCTGATGAGTATTCCTGAGTCTGTGGAGTGGGCTTTATCTG 898
DB 498 GAAAGAAGTGAACCCCTATGAAATTAACCT-GACTGTGTAATGCGGCTTTATCTG 440
QY 899 CCCTGGGCTTCTCGGCTCCATATCATGAGAGACGGCGCTGTGTGCTGACCTGAA 958
DB 439 CCCT-GGCTTCTCGGCTCCATATCATGAGAGACGGCGCTGTGTGCTGACCTGAA 381
QY 959 GTGAGAGCCCAAGCCAGTGAAGAGCTTCTTG-CCCATGAGAGTGAAGAGAG 1017
DB 380 GTGAGAG-MCCAAAGCCAGTGAAGAGCTTCTTGAGAGAGTGAAGAGAG 322
QY 1018 CTTGTGTATGAGAAAGTCCCTTGAGCTTAATCATGTTCCGCAATGAGTCCGCTGT 1077
DB 321 CTTGTGTATGAGAAAGNC-CTTGAGCCCTTAATCATGTTCCGCAATGAGTCCGCTGT 263
QY 1078 ACCTCATCCATGAGTCCCATTTGTACTGTACAGTGAATGAGAGAACTGAATGACGCA 1137
DB 262 ACCTCATCCATG-CCCAATTTGTACTGTACAGTGAATGAGAGAACTGAATGCA-NCA 205
QY 1138 AGGCCCCCTCCGCAACCCCAAGATGAGACAGTCTCTGAGCTGCACTGAGAGGCTT 1197
DB 204 AGGCCCCCTCCGCAACCCCAAGATGAGACAGTCTCTGAGCTGCACTGAGAGGCTT 146
QY 1198 TCTGATATGATGCTGCGCTGCGCCCACTCTCTAGTCTCTGATATTAAGCAATGCTT 1257
DB 145 TCTGATATGATGCTG-CTGCGCCCACTCTCTAG-TCTGATATTAAGCAATGCTT 89
QY 1258 CTTGCTTCTCATGAGGCGCTTGTGACATGAGTGTGCGCTGAGCCCACTCCAGTT 1317
DB 88 CTT-GCTTCTCATGAG-GGNCCTTGTGACATGAGTGTG--CTGNCCTGAGCTCCCACTT 33
QY 1318 TCTCATATTAAGACAGATGCTTCTTCACTT 1349
DB 32 TCTCATATTAAGACAGATGCTTCACTT 1

RESULT 13
US-09-960-352-10323
; Sequence 10323, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10323
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 44-LIB34-041-Q1-E1-C8
US-09-960-352-10323

Query Match 9.4%; Score 298.4; DB 10; Length 420;
Best Local Similarity 81.9%; Pred. No. 3.6e-88;
Matches 344; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 663 AGCGGAGTGGCTGAGATAGTGGCCCATTTGCTCATCATGAGGCGGCTCTTACGCA 722
DB 1 AGCGGAGTGGCTGAGATAGTGGCCCATTTGCTCATCATGAGGCGGCTCTTACGCA 60
QY 723 GTGATGCTGCTTTTCTCTGGAAGAAACTTATCTGTGTCAGAGGACCCAGTATAT 782
DB 61 GTGATGCTGCTTTTCTCTGGAATTAAGACTCTACCTCATCAAGGATACCCAGTATAT 120

QY 783 GTCTCTGACAAAGGAGGCTATACCTAGTAAGCGGTATCCGAGGCGCTGAGAG 842
DB 121 ATTCTCTGACAAAGGAGGCTATACCTCTGTAAGATTAATCCAAAGAGCTGAGAG 180
QY 843 GAATGCGGAGCCCTCATGGAATTAATCTGAGCTGTGAGTGGGCTTTATCTGCT 902
DB 181 GAATTTGGAGGCGCTGATGAGGCTGCTGCTTCAATCTGTGAGATGACGCTTTACTGTCT 240
QY 903 GGCTCTCTCGGCTCATATCATGAGAGACGGCGCTGTGTGCTGAGCTGAGAGTCA 962
DB 241 GAATCTTCTGAGCTCATATCATGAGAGACGGCGCTGTGTGAGCTGAGAGCTTCA 300
QY 963 GAGGCCAAGCCAGTGAAGAGCTTCTTGAGCCCATGAGAGAGTGAAGAGAGCTTG 1022
DB 301 GAGGCTCAAGCCAGTGAAGAGCTTCTTGAGCTTCAATGAGAGAGTGAAGAGAGCTTG 360
QY 1023 TGTATGAGAAAGTCCCTTGAGCTTAACTGATGTTCCGCAATGAGTCCGCTTTGACTTC 1082
DB 361 TGTATGAGAAAGTCTGAGGCGCCCACTGATGTTCTGCAATGAGTGGGCTTTGACTTC 420

RESULT 14
US-09-960-352-7040
; Sequence 7040, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7040
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 30-LIB34-026-Q1-E1-H5
US-09-960-352-7040

Query Match 8.8%; Score 280; DB 10; Length 408;
Best Local Similarity 80.4%; Pred. No. 4.9e-82;
Matches 328; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 767 GGGCACCCAGTATATGCTTCTGCAAAAGGAGGCTTATACCTAGTAAGCGGTATCC 826
DB 1 GGGTACCCAGTATATATTTCTGCAAGGAGGCTTATACCTGTAAGATTAATCC 60
QY 827 GAAGCGGCTGAGAGAGAGTGGGACCCCTCATGAGATTAATCTGAGTCTGTGATGC 886
DB 61 AAAGCAGCTGAGAGAGATTTGGGAGCCCTATGAGGCTCTTCATTTGTGTGATGC 120
QY 887 GGCTTTATCTGCGGCTTCTCTGCTCCATATCATGAGAGACGGCGCTGTGAGT 946
DB 121 AGCTTTATCTGCTGATCTTCTGATCTTCTGATCTTCAATCATGAGAGGCTGTGAG 180
QY 947 GCTGACCTGAAGTCAAGAGCCCAAGCAGTGAAGAGTCTTCTTGGCCCATAGAG 1006
DB 181 GCTGACCTGAAGTCAAGAGCCTCAAGCAGTGAAGAGTCTTCTTGGCCCATAGAG 240
QY 1007 GGTAGAGGAGCTTGTGTATGAGAAAGTCCCTTGAGCTTAACTCATGTTCCGCCAATGG 1066
DB 241 AGTCATATGAGGAGCTGTGTATGAGAGAGTCTGAGGCGCCCACTGATGTTCTGCAATGG 300
QY 1067 TCCGCTTGTACCTCATCATGAGTCCCAATTTGTAAGTCTCAAGTGAAGAGAACT 1126
DB 301 TCTGAGCTTGTACCTCATCATGAGGCGCCCAATCTGTAAGTGAAGAGTGAAGAACT 360
QY 1127 GAATGAGCAGGAGGCTTCCGCAACCCAGATGAGCAAGTCTTCT 1174

RESULT 10
US-09-918-995-30754
; Sequence 30754, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30754
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(488)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30754

Query Match 11.7%; Score 374.2; DB 11; Length 488;
Best Local Similarity 99.2%; Pred. No. 2e-113;
Matches 376; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 503 AGCCATGACACAGAAATGGAGCTGGCCATGGAAACATACCCATGAGCCCTGAGTA 562
DB 110 AGCCATGACACAGAAATGGAGCTGGCCATGGAAACATACCCATGAGCCCTGAGTA 169
QY 563 TATGCGCTGAGCCCATCATCTAGCTTTGTGTGCACTGACCTTGACCAACATGCTGCAC 622
DB 170 TATGCGCTGAGCCCATCATCTAGCTTTGTGTGCACTGACCTTGACCAACATGCTGCAC 229
QY 623 CTATGCTTCAAGTGGAGCCCACTAGTGGGCTGAGACACAGCCGGAGATGGCTGAGTAG 682
DB 230 CTATGCTTCAAGTGGAGCCCACTAGTGGGCTGAGACACAGCCGGAGATGGCTGAGTAG 289
QY 683 CTGAGCCATGCTCATCATGAGTGGCCCAAGGCTCTTCCAGAGTGAATGCTCTTTCTG 742
DB 290 CTGAGCCATGCTCATCATGAGTGGCCCAAGGCTCTTCCAGAGTGAATGCTCTTTCTG 349
QY 743 GGAAGAAAACCTATATCTGTGTCAGGGCACTCAGTATATGCTTTCTGCAAAAGGAGG 802
DB 350 GGAAGAAAACCTATATCTGTGTCAGGGCACTCAGTATATGCTTTCTGCAAAAGGAGG 409
QY 803 CTATACCTAGTAAAGGGTATCCGAAGGGCTGAGAGAAAGTGGGGACCCCTCATG 862
DB 410 CTATACCTAGTAAAGGGTATCCGAAGGGCTGAGAGAAAGTGGGGACCCCTCATG 469
QY 863 GATTATCTGAGCTCTGTG 881
DB 470 GATTATCTGAGCTCTGTG 488

RESULT 11
US-09-918-995-32820
; Sequence 32820, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32820
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(491)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32820

Query Match 10.8%; Score 345.4; DB 11; Length 491;
Best Local Similarity 99.7%; Pred. No. 8.4e-104;
Matches 346; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 503 AGCCATGACACAGAAATGGAGCTGGCCATGGAAACATACCCATGAGCCCTGAGTA 562
DB 145 AGCCATGACACAGAAATGGAGCTGGCCATGGAAACATACCCATGAGCCCTGAGTA 204
QY 563 TATGCGCTGAGCCCATCATCTAGCTTTGTGTGCACTGACCTTGACCAACATGCTGCAC 622
DB 205 TATGCGCTGAGCCCATCATCTAGCTTTGTGTGCACTGACCTTGACCAACATGCTGCAC 264
QY 623 CTATGCTTCAAGTGGAGCCCACTAGTGGGCTGAGACACAGCCGGAGATGGCTGAGTAG 682
DB 265 CTATGCTTCAAGTGGAGCCCACTAGTGGGCTGAGACACAGCCGGAGATGGCTGAGTAG 324
QY 683 CTGAGCCATGCTCATCATGAGTGGCCCAAGGCTCTTCCAGAGTGAATGCTCTTTCTG 742
DB 325 CTGAGCCATGCTCATCATGAGTGGCCCAAGGCTCTTCCAGAGTGAATGCTCTTTCTG 384
QY 743 GGAAGAAAACCTATATCTGTGTCAGGGCACTCAGTATATGCTTTCTGCAAAAGGAGG 802
DB 385 GGAAGAAAACCTATATCTGTGTCAGGGCACTCAGTATATGCTTTCTGCAAAAGGAGG 444
QY 803 CTATACCTAGTAAAGGGTATCCGAAGGGCTGAGAGAAAGTGG 849
DB 445 CTATACCTAGTAAAGGGTATCCGAAGGGCTGAGAGAAAGTGG 491

RESULT 12
US-09-880-107-3166/c
; Sequence 3166, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3166
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 T61801
; NAME/KEY: unsure
; LOCATION: (1)..(504)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3166

Query Match 9.8%; Score 313.4; DB 10; Length 504;
Best Local Similarity 91.4%; Pred. No. 4e-93;

Db	121	CTGGTCTCTGGCCATTGGCCACCCCTTCTCTCCAGCTAGTGGCCATGGGAATGTTGGCTGA	180
Qy	125	AGCGGAGACCAAGCCAGACCAGACGCTGA CTGAACGCTGCTCAGATGGCTGAAGCTTTGA	184
Db	181	AGCGGAGACCAAGCCAGACCAGACGCTGA CTGAACGCTGCTCAGATGGCTGAAGCTTTGA	240
Qy	185	TGCTACCACTCTGGATGACATGGAAACATGCTGTTTTTAAAGGGAGTTTGTGTGAA	244
Db	241	TGCTACCACTCTGGATGACATGGAAACATGCTGTTTTTAAAGGGAGTTTGTGTGAA	300
Qy	245	GATTCACAAATGGGACCGGAGGTTAATCTCAAGAGATGGAAAGATTTTCCACGCTGT	304
Db	301	GATTCACAAATGGGACCGGAGGTTAATCTCAAGAGATGGAAAGATTTTCCACGCTGT	360
Qy	305	GGATGCGCTATTCGGTCAAGGTCAACAAGTGTCTTTCTGATCAAGGGGGACAAAGTCTG	364
Db	361	GGATGCGCTATTCGGTCAAGGTCAACAAGTGTCTTTCTGATCAAGGGGGACAAAGTCTG	420
Qy	365	GGTATATCCCTCTGAATAAGAGGAGAAAGATTAACCAAGTTGCTCCAAGATGAAATTTCC	424
Db	421	GGTATATCCCTCTGAATAAGAGGAGAAAGATTAACCAAGTTGCTCCAAGATGAAATTTCC	480
Qy	425	TGGAATC 431	
Db	481	TGGAATC 487	

SOFTWARE: PatentIn version 3.1
 SEQ ID NO 84
 LENGTH: 1516
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 US-10-175-523-84

Query Match 18.1%; Score 576.8; DB 15; Length 1516;
 Best Local Similarity 67.3%; Pred. No. 1.1e-180;
 Matches 999; Conservative 0; Mismatches 257; Indels 228; Gaps 4;

```

QY 1 CTCTGACGCTCAGCATGCTAGAGGCTAGTCTGGAGACCCGTTGACCTGGGGTTGAGGCC 60
DB 41 CTTTGACGCTGCGCATGCTAGAGACAGTATGAGACATAATATCTGTGATTTGCTGGGCC 100
QY 61 TATGCTGCTCTGCGCATGCTAGGCCACCCCTCTTCTCCGACTAGTCCCATGGAAATTTG 120
DB 101 TGTGCTGCTCTGCGCATGCTAGGCCACCCCTCTTCTCCGACTAGTCCCATGGAAATTTG 160
QY 121 CTGAAAGGCGAAGCCAGACCGACGCTGTAAGCTGTAAGCTGCTGATGGCTGAGACT 180
DB 161 GTGAAATGGAGCCAGACCGACGCTGTAAGCTGTAAGCTGCTGATGGCTGAGACT 220
QY 181 TTGATGCTACACCCCTGATGACATGAAACCATGCTGTTTAAAGGGAGTTGTGT 240
DB 221 TTGACGCTACACCCCTGATGACATGAAACCATGCTGTTTAAAGGGAGTTGTGT 280
QY 241 GGAAGATCACAATGGAGCCGGAGTTTATCTAGAGATGGAAGATTTCCCGAGCC 300
DB 281 GGAAGGATCACAATGGAGCCGGAGTTTATCTAGAGATGGAAGATTTCCCGAGCT 340
QY 301 CTGAGATGCTGATTCCTGCTCAAGGTCAACAGTGTCTTTCTGATCAAGGGAGCAAG 360
DB 341 CAGTGGATGCTGATTCCTGCTCAAGGTCAACAGTGTCTTTCTGATCAAGGGAGCAAG 397
QY 361 TCTGGGATATCCCTCTGAAAGAGAGAAAGATACCAAAAGTTCTCCAGATGAT 420
DB 398 TCTGGGATATCCCTCTGAAAGAGAGAAAGATACCAAAAGTTCTCCAGATGAT 457
QY 421 TTCTGGAATTCCTATCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 458 CTCTGGAATTCCTATCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 517
QY 481 CTGAAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 508
DB 518 GTGAAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 577
QY 509 ----- 508
DB 578 CCAAAAGGAAGCTTCTGCGCTGTGTGGGAATTTGACTGGGCTTGAAGTGGCTTG 637
QY 509 ----- 508
DB 638 AACCTACTACTGCTTCCAGGTAACAAGTTCTGAGATTTAACCCGCTCAGAGAGG 697
QY 509 ----- 508
DB 698 TGCTTCCCAATACCTCTGATGCCGTGACTCTTCTATATCTTCTGCTTGGCAGAGCC 757
QY 509 -----TGACACAGAGATGGAGCTGGCCATGGAGAACAGTACCCACCATGGCCCTGATATA 564
DB 758 ATGCTAAATCTAAAGAAATGGAATCTGCTATGGAATTAAGACCA-----TCTTATGACTT 811
QY 565 TGCGCTGAGCCCACTATGATCTTGTCTGATGATGATGATGATGATGATGATGATGAT 624
DB 812 CGCGTGTAAAGCAGATCTGCGCTGTGTCTGATGATGATGATGATGATGATGATGATGAT 871
QY 625 ATGCTTCAAGTGAAGCCACTACTGCGCTGAGACACAGCGGGGATGGCTGAGATGCT 684
DB 872 ATGCTTCAAGTGAAGCCACTACTGCGCTGAGATGATGATGATGATGATGATGATGATGAT 931
QY 685 GAGCCATTTGCTCATGATGAGCCCAAGGCTCTTCAAGATGATGATGATGATGATGATGAT 744

```

```

DB 932 GGCCCATTCCTCATCATGCGCCCAAGGCTCTTCAAGCATGATGCTGCTTTCTCTGG 991
QY 745 AAGAAACTCTATCTGCTGCTGAGGAGCCAGGATATGCTTCTGTAAGAAAGGAGGCT 804
DB 992 ATGAGAACTCTATCTGATTCAGGAGCATCAAGATATATCTTCTGAGAAAGGAGGCA 1051
QY 805 ATACCTTAAGACGGTTATCCGAAGCGGCTGAGAGAGAGTCCGACCCCTCATGAGGA 864
DB 1052 ATACCTTAAGATGTTATTCGAAGCGGCTGAGAGAGAGTCCGACCCCTCATGAGGA 1111
QY 865 TTATCTGATCTCTGATGAGCGGCTTATCTGCTGAGGCTTCTGAGGCTCATATCA 924
DB 1112 TCAGCTTATATCAATAGATGAGCTTCTTCTGCGCTGCTGCTTCTTCAAGCTTCA 1171
QY 925 TGCAAGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 984
DB 1172 CATGAGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1231
QY 985 AGCTTCTTGGCCCACTGAGAGAGTGAAGAGCTTGTATGAAAGTCCCTTGGCC 1044
DB 1232 AGCTTCTTGGCCCACTGAGAGAGTGAAGAGCTTGTATGAAAGTCCCTTGGCT 1291
QY 1045 CTACTCATGCTCTTCAATGATGCTCCCACTGCTTCTTATCAATGAGGCCAATTAT 1351
DB 1292 CTACTCATGCTCTTCAATGATGCTCCCACTGCTTCTTATCAATGAGGCCAATTAT 1351
QY 1105 GCTACAGTATGAGGAAATGGAATGAGCCAGGCGCTTCCGCAACCCAGAAATGTA 1164
DB 1352 GCTATGAGATGAGGAAATGGAATGAGCCAGGCGCTTCCGCAACCCAGAAATGTA 1411
QY 1165 CCACTCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1221
DB 1412 ACAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1471
QY 1222 CACTCTGATGCTCTGATATTAAGACAGATGCTTCTTCTGCTT 1265
DB 1472 CACTCTGATGCTCTGATATTAAGACAGATGCTTCTTCTGCTT 1515

```

RESULT 7

US-09-918-995-30828
 ; Sequence 30828; Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hysq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FROM VARIOUS CDNA LIBRARIES
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918, 995
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235, 076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 30828
 ; LENGTH: 488
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(488)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-918-995-30828

Query Match 13.4%; Score 427; DB 11; Length 488;
 Best Local Similarity 100.0%; Pred. No. 4.9e-131;
 Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 5 GCAGCTCAGCATGCTAGAGGCTAGTCTGGAGACCCGTTGACCTGGGGTTGAGGCCATG 64
DB 61 GCAGCTCAGCATGCTAGAGGCTAGTCTGGAGACCCGTTGACCTGGGGTTGAGGCCATG 120
QY 65 CTGCTCTGCGCATGCTGCAACCCCTCTCTCTGCACTAGTCCCATGGAATGTTGCTGA 124

```

NUMBER OF SEQ ID NOS: 308
 SOFTWARE: Patent version 3.1
 SEQ ID NO 39
 LENGTH: 1516
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (55) (1437)
 OTHER INFORMATION:
 US-10-316-253-39

Query Match 18.1%; Score 576.8; DB 13; Length 1516;
 Best Local Similarity 67.3%; Pred. No. 1,1e-180;
 Matches 999; Conservative 0; Mismatches 257; Indels 228; Gaps 4;

QY 1 CTCTGACGCTCAGCATGCTGAGGATCTGAGGACACCCGTTGACCTGGGGTTGTGAGCC 60
 DB 41 CTCTGACGCTCAGCATGCTGAGGACAGTATGATCACTAAATATCTGTATGCTGGGCC 100
 QY 61 TATGCTGCTCTGCGCATGCTGACCCCTCTTCTCCGACTAGTGGCCATGGGAATGTTG 120
 DB 101 TGTGCTGCTCTGCGCATGCTGACCCCTCTTCTCCGACTAGTGGCCATGGGAATGTTG 160
 QY 121 CTGAAGGCGAGACCAAGCCAGACCCAGACCTGACCTGAACGCTGCTGATGCTGAGCT 180
 DB 161 GTTAAATGGAACCAAGCCAGACCTGATATTAATCAACCTGCTGATGCTGAGCT 220
 QY 181 TTGATCTACACCTCTGATGACATGGAACCATGCTGTTTTTAAAGGAGTTGTGT 240
 DB 221 TTGACCTACACCATGATGACATGGAACCATGCTGTTTTTAAAGGAGTTGTGT 280
 QY 241 GGAAGGCTCAAAATGGAACCGGAGTATCTCAAGAGATGGAATTTCCCAAGCC 300
 DB 281 GGAAGGCTCAAAATGGAACCGGAGTATCTCAAGAGTATGGAATTTCCCAAGCT 340
 QY 301 CTGTGATGCTGATCTCCGTCAAGTCAACAAGTCTTCTGATCAAGGGGACAAAG 360
 DB 341 CAGTGAATGCTGATCTCCGTCAAGTCAACAAGTCTTCTGATCAAGGGGACAAAG 397
 QY 361 TCTGGTATACCTCTCTGAAAGAGAGAAAGATACCAAAAGTTGCTCAAGATGAT 420
 DB 398 TCTGGTATACCTCTCTGAAAGAGAGAAAGATACCAAAAGTTGCTCAAGATGAT 457
 QY 421 TTCTCTGAAATCCCACTCCCACTGATGACATGGAATGCAACCGTGGAGAAATGTCAG 480
 DB 458 CTCTCTGAAATCCCACTCCCACTGATGACATGGAATGCAACCGTGGAGAAATGTCAG 517
 QY 481 CTGAAGGCGCTCTCTTCTTCAAGGCCA----- 508
 DB 518 GTTAAAGTGTCTTCTTCTTCAAGGTTAACCGCAAGTGTCTTGGGACTTTGCCAAGAA 577
 QY 509 ----- 508
 DB 578 CCAAAAGAAAGCTTCTGCTGCTGTGTGGAATGCACTGGGCTTGAAGTGGCTTG 637
 QY 509 ----- 508
 DB 638 AACGCTACTGCTTCCAGGTTACAAAGTTCTGAGATTTAAACCCGTCACAGAGAGG 697
 QY 509 ----- 508
 DB 698 TGCTCTCCAGATACCTCTGAGATGCCGTGACTATCTTCAATCTGCTCTGGAGAGGCC 757
 QY 509 -----TGACACAGAAATGGAGCTGGCCATGGGAACAGTACCCACATGGCCCTGAGTATA 564
 DB 758 ATGCTAACTGAATGAATGAATGCTCATGGAATGACCA-----TCTATGATTT 811
 QY 565 TGGCTGTAGCCCACTATGATCTTGTCTGCACTGAGCTTGAACACATGCTGCACT 624
 DB 812 CGGCTGTGTAGCCCACTATGATCTTGTCTGCACTGAGCTTGAACACATGCTGCACT 871
 QY 625 ATGCTTGTAGTGGAGCCCACTATGATGCTGTGAGACACAGCGGGAATGCTGCACTGCT 684

DB 872 ATGCTTGTAGTGGAGCCCACTATGATGCTGTGAGCTTCCAGCGGTATGAGTGCATAGCT 931
 QY 685 GGGCCATTTGCTCATGATGAGCCCAAGGTCCTTACAGATGATGCTGCTTTCTCTGGG 744
 DB 932 GGGCCATTTGCTCATGATGAGCCCAAGGTCCTTACAGATGATGCTGCTTTCTCTGGG 991
 QY 745 AAGAAATCTATCTGATGAGGAGGACCAAGGATATGCTTCTGACAAAGAGAGCT 804
 DB 992 ATGAGAAATCTATCTGATGAGGAGGACCTCAAGTATATGCTTCTGACAAAGAGGAGCA 1051
 QY 805 ATACCTTATGAGGAGGATATCCGAGCGGCTGAGAGAGAAAGTCCGACCTCATGGA 864
 DB 1052 ATACCTTATGAGTGTATATCCAAAGCGGCTGAGAGAGAAAGTGTGAGGAGCTCCGGGA 1111
 QY 865 TTATCTGAGCTTGTGATGAGGAGGCTTTATGCTGCTGCTGCTTCTGCTCATATCA 924
 DB 1112 TCAGCTTATGATCAATGATGAGGCTTTCTGCTGCTGCTTCTTCAAGCTTACGTCA 1171
 QY 925 TGAGAGAGGAGGCTGTGTGCTGAGCTGAGCTGAGAGGAGGAGGAGGAGGAGGAG 984
 DB 1172 CATGAGAGGAGGCTTGTGTGCTGAGCTGAGCTGAGAGGAGGAGGAGGAGGAGGAG 1231
 QY 985 AGCTTCTTGGCCCAAGAGGATGAGGAGGCTTGTGTATGAGAAAGTCCCTTGGCC 1044
 DB 1232 AGCTTCTTGGCCCAAGAGGATGAGGAGGCTTGTGTATGAGAAAGTCCCTTGGCT 1291
 QY 1045 CTATCTATGTTCCGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
 DB 1292 CCACTATGATGCTTCTTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1351
 QY 1105 GCTACGATGATGAGGAAATGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1164
 DB 1352 GCTATGAGGATGAGGAAATGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1411
 QY 1165 CCACTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1221
 DB 1412 AAGAGATCTTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1471
 QY 1222 CACTCTTATGTTCTTCAATTAAGACAGATGCTTCTTCTGCTT 1265
 DB 1472 CACTCTTATGTTCTTCAATTAAGACAGATGCTTCTTCTGCTT 1515

RESULT 6
 US-10-175-523-84
 ; Sequence 84, Application US/10175523
 ; Publication No. US20030096264A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Brochman, Jeffrey
 ; APPLICANT: Evans, David
 ; APPLICANT: Hook, Derek
 ; APPLICANT: Klimczak, Leszek
 ; APPLICANT: Laeng, Pascal
 ; APPLICANT: Paley, Michael
 ; APPLICANT: Rajan, Prithi
 ; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
 ; FILE REFERENCE: 3235/10795-US3
 ; CURRENT APPLICATION NUMBER: US/10/175,523
 ; PRIOR FILING DATE: 2002-06-18
 ; PRIOR APPLICATION NUMBER: US 60/299,151
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: US 60/317,828
 ; PRIOR FILING DATE: 2001-09-07
 ; PRIOR APPLICATION NUMBER: US 60/325,150
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US 60/333,047
 ; PRIOR FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: US 60/349,936
 ; PRIOR FILING DATE: 2002-01-18
 ; PRIOR APPLICATION NUMBER: US 60/361,834
 ; PRIOR FILING DATE: 2002-03-04
 ; NUMBER OF SEQ ID NOS: 197

CURRENT APPLICATION NUMBER: US/10/105, 891
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER: 09/668, 117
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/552, 929
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 91
SOFTWARE: pc_Fl_genes Version 2.0
SEQ ID NO 19
LENGTH: 1631
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (71) .. (1459)
US-10-105-891-19

Query Match 35.5%; Score 1129.8; DB 15; Length 1631;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1354; Conservative 0; Mismatches 2; Indels 213; Gaps 1;

1 CTCTGAGCTCAGCATGCTAGGTAAGTGGAGACCCCTTGCACCTGGGTTGTGAGCC 60
DB CTCTGAGCTCAGCATGCTAGGTAAGTGGAGACCCCTTGCACCTGGGTTGTGAGCC 116
QY TATGCTGCTCTGAGCATGCTAGGTAAGTGGAGACCCCTTGCACCTGGGTTGTGAGCC 120
DB TATGCTGCTCTGAGCATGCTAGGTAAGTGGAGACCCCTTGCACCTGGGTTGTGAGCC 176
QY CTGAGGCGGAGACCAAGCCAGACCGTGAAGTGAAGCTGAGTGAAGTGAAGTGAAGT 180
DB CTGAGGCGGAGACCAAGCCAGACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 226
QY TTAGTGTACCAAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 240
DB TTAGTGTACCAAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 296
QY GGAAGATCAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 300
DB GGAAGATCAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 356
QY CTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 360
DB CTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 416
QY TCTGGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 420
DB TCTGGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 476
QY TTTCTGGAATCCATCCCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 480
DB TTTCTGGAATCCATCCCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 536
QY CTGAGGCGCTCTCTCTCTCA----- 502
DB CTGAGGCGCTCTCTCTCTCA----- 502
QY 503 ----- 502
DB 503 ----- 502
QY 597 CCATGAAGAGCTTCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
DB 503 ----- 502
QY 657 GCCGCTACTAGCTTCCAGGGTACCAATTCTGCGCTTCCAGCCCTGTCAGGGGAGAG 716
DB 503 ----- -AGCC 507
QY 717 TGCCTCCAGGTATCCCGGGATGTCAGACTACTTCAATGCTGCTGCTGCTGCTGCTG 776
DB 508 ATGAGACAGGAATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGAT 567
QY 777 ATGAGACAGGAATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGG 836

QY 568 GGTGAGCCCACTAGTCTTGTCTGCACTGACGTCTGCAACCAATGAGCCACTATG 627
DB GGTGAGCCCACTAGTCTTGTCTGCACTGACGTCTGCAACCAATGAGCCACTATG 836
QY 628 CTTTCAATGAGGACCACTAGTCTTGTCTGCACTGACGTCTGCAACCAATGAGCC 687
DB CTTTCAATGAGGACCACTAGTCTTGTCTGCACTGACGTCTGCAACCAATGAGCC 956
QY 688 CCAATGCTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 747
DB CCAATGCTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1016
QY 748 AAAAATCTATCTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 807
DB AAAAATCTATCTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1076
QY 808 CCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
DB CCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1136
QY 868 TCTTGAAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 927
DB TCTTGAAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1196
QY 928 CAGAGCGGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 987
DB CAGAGCGGCGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1256
QY 988 TTTCTTGGCCCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1047
DB TTTCTTGGCCCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1316
QY 1048 ACTCATGTTCCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1107
DB ACTCATGTTCCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1376
QY 1317 ACTCATGTTCCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1376
DB ACTCATGTTCCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1436
QY 1437 ACTCATGTTCCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1436
DB ACTCATGTTCCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
QY 1497 CTAGTCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1556
DB CTAGTCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1616
QY 1288 TGAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1347
DB TGAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1616
QY 1348 TGAATGCA 1356
DB 1617 TGAATGCA 1625

RESULT 5
US-10-316-253-39
Sequence 39, Application US/10316253
Publication No. US20030162706A1
GENERAL INFORMATION:
APPLICANT: The Procter & Gamble Company
APPLICANT: Peters, Kevin
APPLICANT: Thompson, Larry
APPLICANT: Wang, Feng
APPLICANT: Greis, Kenneth
TITLE OF INVENTION: Angiogenesis Modulating Proteins
FILE REFERENCE: 885M
CURRENT APPLICATION NUMBER: US/10/316, 253
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 60/355, 295
PRIOR FILING DATE: 2002-02-08

PRIOR FILING DATE: 2000-09-22
 PRIOR APPLICATION NUMBER: 09/552,929
 PRIOR FILING DATE: 2000-04-18
 NUMBER OF SEQ ID NOS: 91
 SOFTWARE: pc_Fl_genes Version 2.0
 SEQ ID NO 19
 LENGTH: 1631
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (71) ..(1459)
 US-10-125-237-19

Query Match 35.5%; Score 1129.8; DB 15; Length 1631;
 Best Local Similarity 86.3%; Pred. No. 0;
 Matches 1354; Conservative 0; Mismatches 2; Indels 213; Gaps 1;

QY 1 CTCCTGAGCTCAGCATGGCTAGGGTACTGGAGGACCCGTTGCACTGGGGTTGTGAGCC 60
 DB 57 CTCCTGAGCTCAGCATGGCTAGGGTACTGGAGGACCCGTTGCACTGGGGTTGTGAGCC 116
 QY 61 TATGCTGCTCTGGCCATTGCCACCCCTCTTCTCCGACTAGTGCCCATGGGAATGTTG 120
 DB 117 TATGCTGCTCTGGCCATTGCCACCCCTCTTCTCCGACTAGTGCCCATGGGAATGTTG 176
 QY 121 CTGAAGCGGAGCAAGCCAGACCCAGACGTTGACCTGAACGCTGCTCAGATGGCTGAGCT 180
 DB 177 CTGAAGCGGAGCAAGCCAGACCCAGACGTTGACCTGAACGCTGCTCAGATGGCTGAGCT 236
 QY 181 TTGATGCTACCACTCTGATGACATGGAACATGCTGTTTAAAGGGAATGTTGT 240
 DB 237 TTGATGCTACCACTCTGATGACATGGAACATGCTGTTTAAAGGGAATGTTGT 296
 QY 241 GGAAGAGTCAAAATGGAGCCGGAGTTAATCTCAGAGATGGAGAATTTCCCAAGCC 300
 DB 297 GGAAGAGTCAAAATGGAGCCGGAGTTAATCTCAGAGATGGAGAATTTCCCAAGCC 356
 QY 301 CTGTGATGCTGATTCCTGCAAGGTGACAAAGTGTCTTCTGATGCAAGGGGAGCAAG 360
 DB 357 CTGTGATGCTGATTCCTGCAAGGTGACAAAGTGTCTTCTGATGCAAGGGGAGCAAG 416
 QY 361 TCTGGATATACCTCTGAAAGAGGAGGATATCCCAAGTTGCTCCAGATGAT 420
 DB 417 TCTGGATATACCTCTGAAAGAGGAGGATATCCCAAGTTGCTCCAGATGAT 476
 QY 421 TTCTGGAATCCATCCCACTGATGCACTGTGATGTCACCGTGAAGATGTCAAG 480
 DB 477 TTCTGGAATCCATCCCACTGATGCACTGTGATGTCACCGTGAAGATGTCAAG 536
 QY 481 CTGAAGCGCTCTCTTCTTCCA----- 502
 DB 537 CTGAAGCGCTCTCTTCTTCCAAGGTGACCGGAGTGTCTGGAATTGGAACCGGAA 596
 QY 503 ----- 502
 DB 597 CCATGAAGAGCGTCTCTGCGCAGCTGTGGGAATGCTCTCTGCTCCCTGAGATGCTGG 656
 QY 503 ----- 502
 DB 657 GCCGCTACTACTGCTTCCAGGGTAAACCAATTCTGCGCTTGAACCTGTGAGGGAGAGG 716
 QY 503 -----AGGCC 507
 DB 717 TGCCTCCAGGTACCCCGGGATGTCGAGACTACTTCATGCTCTGCTCCCTGCGAGAGGCC 776
 QY 508 ATGAACACAGGAATGGAGCTGGCCATGGGAAAGTACCAACCAATGGCCCTGAGATATGTC 567
 DB 777 ATGAACACAGGAATGGAGCTGGCCATGGGAAAGTACCAACCAATGGCCCTGAGATATGTC 836
 QY 568 GCTGTAGCCCACTAGTCTGTCTGCACTGACGTGTGAACAACATGATGCCACTATG 627
 DB 837 GCTGTAGCCCACTAGTCTGTCTGCACTGACGTGTGAACAACATGATGCCACTATG 896

QY 628 CCTTCAATGGGAGCCACTACTGGGCTGTGGAGACAGCCGGAGATGGCTGATAGTGGC 687
 DB 897 CCTTCAATGGGAGCCACTACTGGGCTGTGGAGACAGCCGGAGATGGCTGATAGTGGC 956
 QY 688 CCATTCCTCATAGTGGGCCCAAGGCTCTTCAAGCAGTGAATGCTGCTTCTCTGGGAG 747
 DB 957 CCATTCCTCATAGTGGGCCCAAGGCTCTTCAAGCAGTGAATGCTGCTTCTCTGGGAG 1016
 QY 748 AAAAATCTATCTGTGTCAGAGGACCCAGGATATATGCTTCTGACAAAGGAGGCTATA 807
 DB 1017 AAAAATCTATCTGTGTCAGAGGACCCAGGATATATGCTTCTGACAAAGGAGGCTATA 1076
 QY 808 CCTTAGTAAGCGGTTATCCGAGAGCGCTGGAGAGGAGGAGGAGCCCTCATAGGATTA 867
 DB 1077 CCTTAGTAAGCGGTTATCCGAGAGCGCTGGAGAGGAGGAGGAGGAGCCCTCATAGGATTA 1136
 QY 868 TCTGGAATCTGTGATGACGCGCTTATATCTGCTGCTGCTTCTGCTCATATCATG 927
 DB 1137 TCTGGAATCTGTGATGACGCGCTTATATCTGCTGCTGCTTCTGCTCATATCATG 1196
 QY 928 CAGAGCGGCGCTGTGTGCTGAGCTGAGCTTGAATCAGAGAGCCCAAGCCATGAGCAGAGC 987
 DB 1197 CAGAGCGGCGCTGTGTGCTGAGCTGAGCTTGAATCAGAGAGCCCAAGCCATGAGCAGAGC 1256
 QY 988 TTCTTGGCCCATGAGAGGTAGAGGAGCTGTGTATGGAAGTCCCTTGGCCCTTA 1047
 DB 1257 TTCTTGGCCCATGAGAGGTAGAGGAGCTGTGTATGGAAGTCCCTTGGCCCTTA 1316
 QY 1048 ACTCATGTTCCGCAATGATGTCGCGCTTATCTGACTCATGATGATGCTTCCCAATTTGACTGCT 1107
 DB 1317 ACTCATGTTCCGCAATGATGTCGCGCTTATCTGACTCATGATGATGCTTCCCAATTTGACTGCT 1376
 QY 1108 ACAGTATGTGAGAAATGATGACAGCCAGGCTTCCGCAACCCAGATGTGACCA 1167
 DB 1377 ACAGTATGTGAGAAATGATGACAGCCAGGCTTCCGCAACCCAGATGTGACCA 1436
 QY 1168 GTCTCGGCGCTGACATCTGAGGGGCGCTTGAATGATGCTGAGCCCAAGCTC 1227
 DB 1437 GTCTCGGCGCTGACATCTGAGGGGCGCTTGAATGATGCTGAGCCCAAGCTC 1496
 QY 1228 CTAGTTCCTATATTAAGCAGATGCTTCTGCTTCTGCTGAGGGGCTTCTGACA 1287
 DB 1497 CTAGTTCCTATATTAAGCAGATGCTTCTGCTTCTGCTGAGGGGCTTCTGACA 1556
 QY 1288 TGAATCTGCGCTGGCCCACTCCCAAGTTCTCATATTAAGCAGATGCTTCTTAC 1347
 DB 1557 TGAATCTGCGCTGGCCCACTCCCAAGTTCTCATATTAAGCAGATGCTTCTTAC 1616
 QY 1348 TTGAATCAA 1356
 DB 1617 TTGAAAAAA 1625

RESULT 4

US-10-105-891-19
 Sequence 19, Application US/10105891
 Publication No. US20030073099A1
 GENERAL INFORMATION:
 APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Zhou, Ping
 APPLICANT: Xu, Feiyan
 APPLICANT: Zhou, Qiang A.
 APPLICANT: Zhang, Jie
 APPLICANT: Wehrman, Tom
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Dimaanac, Radoje T.
 TITLE OF INVENTION: No. US20030073099A1el Nucleic Acids and
 FILE REFERENCE: 791CIP2A

Db 11888 TGGTTCTTCACTGGAATCAAGGAGCCTTGGTCCGTGAACAATCTTCTTTCTTGAATGGA 11947
Qy 1397 AAAATTAGCACTTCTCTCTTTGAGGGTGTGAGCTTCAACAGAGCTGTGAGAAACAAGGA 1456
Db 11948 AAAGTTAGCACTTCTCTCTTTGAGGGTGTGAGCTTCAACAGAGCTGTGAGAAACAAGGA 12007
Qy 1457 GGGGAGCACTAAAGGGGCAAACTATCTGTGCGAGATGATTTCTTAGTCCAGATCAATAA 1516
Db 12008 GGGGAGCACTAAAGGGGCAAACTATCTGTGCGAGATGATTTCTTAGTCCAGATCAATAA 12067
Qy 1517 CTAGCTCTTTGCAAGATCTATACATAGTGGGGGAAAGAGACCAAGATCCGAAGAG 1576
Db 12068 CTAGCTCTTTGCAAGATCTATACATAGTGGGGGAAAGAGACCAAGATCCGAAGAG 12127
Qy 1577 AACAGCTGAGTTTATACAGCAAGTAAAGGTGAGCTAGAGCTCTGATTTCACTTGTG 1636
Db 12128 AACAGCTGAGTTTATACAGCAAGTAAAGGTGAGCTAGAGCTCTGATTTCACTTGTG 12187
Qy 1637 TAGATGGCCACAAACCGGCAAGGATGAGAAACAAGGGCTGGGGCACTATGCA 1696
Db 12188 TAGATGGCCACAAACCGGCAAGGATGAGAAACAAGGGCTGGGGCACTATGCA 12247
Qy 1697 TGTGCAAGAGATTTGCTCAGAGTTGTGGGGTGAAGAGTCAATCTGGGGACTCAAA 1756
Db 12248 TGTGCAAGAGATTTGCTCAGAGTTGTGGGGTGAAGAGTCAATCTGGGGACTCAAA 12307
Qy 1757 TTAATGTTCTGGGTGATTCAGATTAACAACATAGGCTTGTGTGCAATGATTAAGCA 1816
Db 12308 TTAATGTTCTGGGTGATTCAGATTAACAACATAGGCTTGTGTGCAATGATTAAGCA 12367
Qy 1817 TGACAAGTGAATGAAGTTGAAGTGGGGAAACAGAAATACACAGCTGTGTGCAAGG 1876
Db 12368 TGACAAGTGAATGAAGTTGAAGTGGGGAAACAGAAATACACAGCTGTGTGCAAGG 12427
Qy 1877 AAAGCTGAGAGAGAGAGAAAGATGAATGCAACATGAGACCAATTTGAGAAACAAGT 1936
Db 12428 AAAGCTGAGAGAGAGAGAAAGATGAATGCAACATGAGACCAATTTGAGAAACAAGT 12487
Qy 1937 CCTGGGAGATCTTGTCTGAGAGCTCTCAGAGACTTTGTCTGAGACAGAGATCTGGCTTA 1996
Db 12488 CCTGGGAGATCTTGTCTGAGAGCTCTCAGAGACTTTGTCTGAGACAGAGATCTGGCTTA 12547
Qy 1997 ATTAGCTCTGGGTGATTCGACATCTAGACCAAGAAATGTAATGCGCAAGAGAGTGG 2056
Db 12548 ATTAGCTCTGGGTGATTCGACATCTAGACCAAGAAATGTAATGCGCAAGAGAGTGG 12607
Qy 2057 AGTGTGTTGATTTCTGGTCAATAACTGTGAAGGCTTTGGGTTCCAATTTGCTGAC 2116
Db 12608 AGTGTGTTGATTTCTGGTCAATAACTGTGAAGGCTTTGGGTTCCAATTTGCTGAC 12667
Qy 2117 AGCTGTTTAACTGGGAATTTGGGCTTAGACTATAGTAGTATGTCTCAGACAAAGCCCTA 2176
Db 12668 AGCTGTTTAACTGGGAATTTGGGCTTAGACTATAGTAGTATGTCTCAGACAAAGCCCTA 12727
Qy 2177 TTCTCTCAGCTGCTTTTCAACCAAGCTGAGGCTTGGAGGCTGTGTTTCAAGCTCAAAA 2236
Db 12728 TTCTCTCAGCTGCTTTTCAACCAAGCTGAGGCTTGGAGGCTGTGTTTCAAGCTCAAAA 12787
Qy 2237 AATAGCTGAGTTTCCAGAGAGGCTTATTTCTAGGCTTCCGTGCTCAGGCTCAATT 2296
Db 12788 AATAGCTGAGTTTCCAGAGAGGCTTATTTCTAGGCTTCCGTGCTCAGGCTCAATT 12847
Qy 2297 TCTCTTCTGTTAAATAGACAAATGCAACCTCCAGCTTCCAGTGAATATAGATCTC 2356
Db 12848 TCTCTTCTGTTAAATAGACAAATGCAACCTCCAGCTTCCAGTGAATATAGATCTC 12907
Qy 2357 AAACCATCCCTTGAAGCTGTTGGGAAGGGGCTTGAAGTGAACCCAGACGTGGCTC 2416
Db 12908 AAACCATCCCTTGAAGCTGTTGGGAAGGGGCTTGAAGTGAACCCAGACGTGGCTC 12967
Qy 2417 ATGGCTCATGTGATGTGAGTCAAGCCCTCCCAACTGTCAAGCAATTTGCTCCGTAGA 2476
|||||

Db 12968 ATGGCTCATGTGATGTGAGTCAAGCCCTCCCAACTGTCAAGCAATTTGCTCCGTAGA 13027
Qy 2477 CTTTGAATGGGTAGAGTATGCTTAACAGCTCTGATGTCAACAAGGCTTTGACTGGG 2536
Db 13028 CTTTGAATGGGTAGAGTATGCTTAACAGCTCTGATGTCAACAAGGCTTTGACTGGG 13087
Qy 2537 AGGCCAGCTATAGATGCTCAGGCTTAAAGGCTGGAGCTGGGGGACAGTGTCTCAG 2596
Db 13088 AGGCCAGCTATAGATGCTCAGGCTTAAAGGCTGGAGCTGGGGGACAGTGTCTCAG 13147
Qy 2597 ATTAGGCTTAACTAGAAATTTGACTGAGCTGAGAAACAAGATTAAGGGCCACAGCA 2656
Db 13148 ATTAGGCTTAACTAGAAATTTGACTGAGCTGAGAAACAAGATTAAGGGCCACAGCA 13207
Qy 2657 GGGTGTGGGTCTACTCTCTTGAAGACCTTGAAGCTTACTTTTCAATTTCTTAATGTTGTC 2716
Db 13208 GGGTGTGGGTCTACTCTCTTGAAGACCTTGAAGCTTACTTTTCAATTTCTTAATGTTGTC 13267
Qy 2717 TTGATGCTACCCCTCAAGGGTGGCTGCTAGTCTTAAGGGGTGAGACAAAGACAGAT 2776
Db 13268 TTGATGCTACCCCTCAAGGGTGGCTGCTAGTCTTAAGGGGTGAGACAAAGACAGAT 13327
Qy 2777 TTCAAGTCTGATCTTATCAAGTTCAATGCACTAACAATTTGGGACACTGTGCAATGCGC 2836
Db 13328 TTCAAGTCTGATCTTATCAAGTTCAATGCACTAACAATTTGGGACACTGTGCAATGCGC 13387
Qy 2837 AGGAGCTTAGAGGTGTCTAAACAGTTATCCAAACAATGTGATACCAAGTTAACTTTC 2896
Db 13388 AGGAGCTTAGAGGTGTCTAAACAGTTATCCAAACAATGTGATACCAAGTTAACTTTC 13447
Qy 2897 TCTTGTTCAGAGGAGGAGATTAAGTCTACCTTCTCTCTCTCCACAGTGTTC 2956
Db 13448 TCTTGTTCAGAGGAGGAGATTAAGTCTACCTTCTCTCTCTCCACAGTGTTC 13507
Qy 2957 TCTTGAAGGAATTCCTTAGTCTTCTCAAGGAACTCCAGAAATGTTTGTTCAGTC 3016
Db 13508 TCTTGAAGGAATTCCTTAGTCTTCTCAAGGAACTCCAGAAATGTTTGTTCAGTC 13567
Qy 3017 AGTTTAGGCTGTATTAAGATATCTTGAAGTGGGTATCTATCAGCAATAGAAATTTA 3076
Db 13568 AGTTTAGGCTGTATTAAGATATCTTGAAGTGGGTATCTATCAGCAATAGAAATTTA 13627
Qy 3077 TTGTTCACAATTCGAGAGCTGGAATAATCAAGATCAAGGCTCCAGAGGTTCAAGTGTCT 3136
Db 13628 TTGTTCACAATTCGAGAGCTGGAATAATCAAGATCAAGGCTCCAGAGGTTCAAGTGTCT 13687
Qy 3137 GCTGAGTCTTGTCTGCTTCAAGATGCAACCTTTTGTCTGTCTCA 3186
Db 13688 GCTGAGTCTTGTCTGCTTCAAGATGCAACCTTTTGTCTGTCTCA 13737
|||||

RESULT 3
US-10-125-237-19
; Sequence 19, Application US/10125237
; Publication No. US20030022329A1
GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyun
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aiding J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030022329A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2ADIV
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
;


```

QY 2401 ACCGAGACTGGCTCATGGCTCATGTGATGTGAGTACAGCCCTCCCACTGTGACG 2460
DB 2401 ACCGAGACTGGCTCATGGCTCATGTGATGTGAGTACAGCCCTCCCACTGTGACG 2460
QY 2461 CATTGCTCCGTGAGACTTTGATGGGTAGAGTACAGTCAAGAGCTTGTGACAC 2520
DB 2461 CATTGCTCCGTGAGACTTTGATGGGTAGAGTACAGTCAAGAGCTTGTGACAC 2520
QY 2521 AAGGCTTTTACCTGGAGGCGAGGCTATAGAGTGTCCAGCTTAAAGGCTGGAGCTG 2580
DB 2521 AAGGCTTTTACCTGGAGGCGAGGCTATAGAGTGTCCAGCTTAAAGGCTGGAGCTG 2580
QY 2581 GGGGACAGTGTCTCAGATTAGAGTCTTAAGTGAAGTTGACTGAGCTGAGACAGAGT 2640
DB 2581 GGGGACAGTGTCTCAGATTAGAGTCTTAAGTGAAGTTGACTGAGCTGAGACAGAGT 2640
QY 2641 TAGGGGCGAAGCAGAGGGTGTGGGTCTAATCTCTTAAAGACCTTGAAGCTTTT 2700
DB 2641 TAGGGGCGAAGCAGAGGGTGTGGGTCTAATCTCTTAAAGACCTTGAAGCTTTT 2700
QY 2701 CATTCCTAATGGTGTGGTGGTGAAGCTTACCTCAAGGGTGGTGGTGTGCTAAGGGGTG 2760
DB 2701 CATTCCTAATGGTGTGGTGGTGAAGCTTACCTCAAGGGTGGTGGTGTGCTAAGGGGTG 2760
QY 2761 GAGACAAAGACAGAGTTCAGGTCTGTCTTATCAAGTTCAATGACTACACTTGGAGCC 2820
DB 2761 GAGACAAAGACAGAGTTCAGGTCTGTCTTATCAAGTTCAATGACTACACTTGGAGCC 2820
QY 2821 ACTGTGATCATGTCAGGAGGAGCTTAAGAGTGTCTTAAAGATTATCAACACTGTGATA 2880
DB 2821 ACTGTGATCATGTCAGGAGGAGCTTAAGAGTGTCTTAAAGATTATCAACACTGTGATA 2880
QY 2881 CCCAAGTTAATCTTCTCTTGTGTTTCAAGGAGGAGTAAAGTCTCCCTTCTCTCT 2940
DB 2881 CCCAAGTTAATCTTCTCTTGTGTTTCAAGGAGGAGTAAAGTCTCCCTTCTCTCT 2940
QY 2941 TTCTCTCCAGTGTCTCTTTCAGGAGAACTCTTACCTGTCTCCAGGAGACTCCAGAA 3000
DB 2941 TTCTCTCCAGTGTCTCTTTCAGGAGAACTCTTACCTGTCTCTCCAGGAGACTCCAGAA 3000
QY 3001 ATGTTTGTTCAGTCAAGTTTAAAGTGTCTTAAAGAAATATCTTAAAGTGGTAACTAT 3060
DB 3001 ATGTTTGTTCAGTCAAGTTTAAAGTGTCTTAAAGAAATATCTTAAAGTGGTAACTAT 3060
QY 3061 CAGCAATAGAAATTTATTTGTTCATTTCTGAGGCTGAGAAATCCAAGATCAAGGCTCC 3120
DB 3061 CAGCAATAGAAATTTATTTGTTCATTTCTGAGGCTGAGAAATCCAAGATCAAGGCTCC 3120
QY 3121 ACGAGGTTCAAGTGTCTGAGAGTGTGTTCTTCTGAAAGATGAGACCTTTTGTCTGTG 3180
DB 3121 ACGAGGTTCAAGTGTCTGAGAGTGTGTTCTTCTGAAAGATGAGACCTTTTGTCTGTG 3180
QY 3181 TTCTCA 3186
DB 3181 TTCTCA 3186

```

```

RESULT 2
US-09-900-448-3
; Sequence 3, Application US/0900448
; Publication No. US20030220488A1
; GENERAL INFORMATION:
; APPLICANT: CERCARDI, Toni et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001272
; CURRENT APPLICATION NUMBER: US/09/900,448
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13737

```

```

; TYPE: DNA
; ORGANISM: Human
US-09-900-448-3
Query Match 71.5%; Score 2279.2; DB 13; Length 13737;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 2554; Conservative 0; Mismatches 3; Indels 253; Gaps 2;
QY 630 TTCAGTGGGAGCCACTACTGGGTCTGAGACACAGCCGGGATGGGATGCTGGCC 689
DB 10928 TCCACAGGAGCCACTACTGGGTCTGAGACACAGCCGGGATGGGATGCTGGCC 10987
QY 690 ATTGCTCATAGTGGGCCAGGGTCTTTCAGAGATGATGCTCTTTCTGGAGAA 749
DB 10988 ATTGCTCATAGTGGGCCAGGGTCTTTCAGAGATGATGCTCTTTCTGGAGAA 11047
QY 750 AAACCTATCTGG----- 762
DB 11048 AAACCTATCTGGTCCAGGTGTATTTGGGGAGAGGCTTGAAGTGAAGACTGGACAA 11107
QY 763 -----TCCAGGGCACCCAGTATATGCT 786
DB 11108 CATATCAACTCTGATTTATTTATTAACATCTTGTCTCCAGGACCCAGTATATGCT 11167
QY 787 TCCGACAAAGGAGGCTATACCTTAGTAAAGGCTTATCCGAAGCGGCTGAGAAAG 846
DB 11168 TCCGACAAAGGAGGCTATACCTTAGTAAAGGCTTATCCGAAGCGGCTGAGAAAG 11227
QY 847 TGGGAGCCCTCAGAGATTTATCTGAGCTGTGTGATGCGGCTTTATCTGCTGGGT 906
DB 11228 TGGGAGCCCTCAGAGATTTATCTGAGCTGTGTGATGCGGCTTTATCTGCTGGGT 11287
QY 907 CTTTCGGGTCCATATCAGG----- 927
DB 11288 CTTTCGGGTCCATATCAGGAGGTGAGGGCTTCTGGGTCTTAAAGGACCTTGT 11347
QY 928 ----- 927
DB 11348 CTGTACTGTCTGTGGCATATATCCCAACAGGGCATGAGAAAGGCTAGTCAGATCC 11407
QY 928 ----- 927
DB 11408 CCAGGCGATGAGAGGCTTAGGTCAAGATCCCATGACATGGAAGCCATGATGTTGG 11467
QY 928 -----CAGACGGGCGGTGTGTGGCTGAGACCTGAACTGAGAGCCCAAGCCAC 976
DB 11468 TGCCTTCTCCAGAGAGGCGGTGTGTGGCTGAGACCTGAACTGAGAGCCCAAGCCAC 11527
QY 977 GTGACAGAGCTTCTTGGGCCCATGAGAGGTAGACGAGCTTGTGTATGGAAGATC 1036
DB 11528 GTGACAGAGCTTCTTGGGCCCATGAGAGGTAGACGAGCTTGTGTATGGAAGATC 11587
QY 1037 CCTTGGCCCTTAACTCATGTTCCGCAATGTCCTGGCTTGTATCTCATCATGTCCTCA 1096
DB 11588 CCTTGGCCCTTAACTCATGTTCCGCAATGTCCTGGCTTGTATCTCATCATGTCCTCA 11647
QY 1097 TTTGTACTGCTACAGTATGTGTGAGAAATGATGAGCAAGCCCTTCCCAACCCCA 1156
DB 11648 TTTGTACTGCTACAGTATGTGTGAGAAATGATGAGCAAGCCCTTCCCAACCCCA 11707
QY 1157 GAATGACCAAGTCTCTGGGCTGACACTCAAGAGGGGCTTCTGACATGATGTGGCT 1216
DB 11708 GAATGACCAAGTCTCTGGGCTGACACTCAAGAGGGGCTTCTGACATGATGTGGCT 11767
QY 1217 GGCCCACTCTTACTTCTCTAATAAAGACAGATGCTTCTGCTTCTGACAGAGG 1276
DB 11768 GGCCCACTCTTACTTCTCTAATAAAGACAGATGCTTCTGCTTCTGACAGAGG 11827
QY 1277 GCTTCTGACATGAGTCTGGCTGGCCCACTCCCAAGTTTCTCATATAAAGACAGAT 1336
DB 11828 GCTTCTGACATGAGTCTGGCTGGCCCACTCCCAAGTTTCTCATATAAAGACAGAT 11887
QY 1337 TGGTCTTCACTTGAATCAAGGAGCCTTGTGTGAAACATCTTCTTCTTGAAGTTGA 1396

```


Db 181 TTGATGCTACACACCTTGATGACATGATGAAACATGCTTTTAAAGGGAGTTTGTG 240
Qy 241 GGAAGGTCACAATGGGACCGGAGTTAATCTCAGAGATGGAAGATTTCCCAAGCC 300
Db 241 GGAAGGTCACAATGGGACCGGAGTTAATCTCAGAGATGGAAGATTTCCCAAGCC 300
Qy 301 CTGTGATGCTGATTCGGTCAAGGTCAACAGTGTCTTTCATCAAGGGGACAAAG 360
Db 301 CTGTGATGCTGATTCGGTCAAGGTCAACAGTGTCTTTCATCAAGGGGACAAAG 360
Qy 361 TCTGGTATACCTCTCTGAAAAGAGAGAAAGATACCCAAAGTTCTTCAAGATGAAT 420
Db 361 TCTGGTATACCTCTCTGAAAAGAGAGAAAGATACCCAAAGTTCTTCAAGATGAAT 420
Qy 421 TTCTGGATCCCATCCCACTGAGTGCAGCTGTGAAATGCACCGTGGAAATGTCAG 480
Db 421 TTCTGGATCCCATCCCACTGAGTGCAGCTGTGAAATGCACCGTGGAAATGTCAG 480
Qy 481 CTGAAGGCGTCTCTTCTTCAAGGCGATGGAACAGAGAAATGGGACTGGCAATGGGAACA 540
Db 481 CTGAAGGCGTCTCTTCTTCAAGGCGATGGAACAGAGAAATGGGACTGGCAATGGGAACA 540
Qy 541 GTACCCACATGAGCCCTGAGTATATGCGTGTAGCCCACTTATGTCTTGTCTGACTGA 600
Db 541 GTACCCACATGAGCCCTGAGTATATGCGTGTAGCCCACTTATGTCTTGTCTGACTGA 600
Qy 601 CGTCTGACACATGCTGCACTTATGCTTCACTGAGGACCCCACTACTGCGTCTGACA 660
Db 601 CGTCTGACACATGCTGCACTTATGCTTCACTGAGGACCCCACTACTGCGTCTGACA 660
Qy 661 CCAGCCGGAGTGGCTGGCATAGCTGGCCCATGCTCATGATGGCCCAAGGCTTTCAG 720
Db 661 CCAGCCGGAGTGGCTGGCATAGCTGGCCCATGCTCATGATGGCCCAAGGCTTTCAG 720
Qy 721 CAGTGAATGCTGCTTTTCTCTGGAGAAAGAAACTTATCTGTCTCAAGGACCAAGTAT 780
Db 721 CAGTGAATGCTGCTTTTCTCTGGAGAAAGAAACTTATCTGTCTCAAGGACCAAGTAT 780
Qy 781 ATGTCTTCTGACAAAGGAGGCTATACCTTATAGCGGTTATCCGAGCGCTGGAGA 840
Db 781 ATGTCTTCTGACAAAGGAGGCTATACCTTATAGCGGTTATCCGAGCGCTGGAGA 840
Qy 841 AGGAAGTCGGGAGCCCTCATGGAGATTAATCTGAGCTCTGAGATGCGGCTTATCTGCC 900
Db 841 AGGAAGTCGGGAGCCCTCATGGAGATTAATCTGAGCTCTGAGATGCGGCTTATCTGCC 900
Qy 901 CTGGGCTTCTCGGCTCATATATGAGCAAGGACGCGGCTGTGGTGGCTGGAAGT 960
Db 901 CTGGGCTTCTCGGCTCATATATGAGCAAGGACGCGGCTGTGGTGGCTGGAAGT 960
Qy 961 CAGGAGCCCAAGCGATGAGCAGAGCTTCTTGGCCCATGAGAAAGTGAAGGAGCT 1020
Db 961 CAGGAGCCCAAGCGATGAGCAGAGCTTCTTGGCCCATGAGAAAGTGAAGGAGCT 1020
Qy 1021 TGTGTAATGGAAGTCCCTTGGCCCTTAATCATGTTCCGCAATGCTGCGGCTTGAAC 1080
Db 1021 TGTGTAATGGAAGTCCCTTGGCCCTTAATCATGTTCCGCAATGCTGCGGCTTGAAC 1080
Qy 1081 TCAATCAATGATGATTTGTACTGTACAGTATGTGAGAAACTGAATGACCAAG 1140
Db 1081 TCAATCAATGATGATTTGTACTGTACAGTATGTGAGAAACTGAATGACCAAG 1140
Qy 1141 CCGTTCGCAACCCCAAGATGTAACAGTCTTCTGGGCTGCACTCACTGAGGGGCTTCT 1200
Db 1141 CCGTTCGCAACCCCAAGATGTAACAGTCTTCTGGGCTGCACTCACTGAGGGGCTTCT 1200
Qy 1201 GACATGATGCTGGGCTGGGCGCCCACTCTGATTTCTCTATAATAAGCAATGCTTCT 1260
Db 1201 GACATGATGCTGGGCTGGGCGCCCACTCTGATTTCTCTATAATAAGCAATGCTTCT 1260
Qy 1261 CGCTTCTCATGAGGGGCTTGTGACATGAGTCTGGCTGGCCCACTCTCCCAAGTTCT 1320
Db 1261 CGCTTCTCATGAGGGGCTTGTGACATGAGTCTGGCTGGCCCACTCTCCCAAGTTCT 1320

Qy 1321 CATATAAAGAGATGCTTCTTCACTTGAATCAAGGAGCCTTGGTGGGAAACAATCT 1380
Db 1321 CATATAAAGAGATGCTTCTTCTTCACTTGAATCAAGGAGCCTTGGTGGGAAACAATCT 1380
Qy 1381 TCTTTCTTGAATGAAAAGTTAGCACTTCTCTTGAAGGAGTGTGAGCTCAACAAGGC 1440
Db 1381 TCTTTCTTGAATGAAAAGTTAGCACTTCTCTTGAAGGAGTGTGAGCTCAACAAGGC 1440
Qy 1441 TGTGAGAAACAAGGAGGAGCACTAAGGGCAACCTATCTCTGCGCATGATGATTTCT 1500
Db 1441 TGTGAGAAACAAGGAGGAGCACTAAGGGCAACCTATCTCTGCGCATGATGATTTCT 1500
Qy 1501 AGGTCCAGATCAATACTAGCTTCTTGAAGCACTATCTTCACTAATGAGGGGAGAAAGAA 1560
Db 1501 AGGTCCAGATCAATACTAGCTTCTTGAAGCACTATCTTCACTAATGAGGGGAGAAAGAA 1560
Qy 1561 CCAGAGTCGGAAGAGAAAGCTGAGTTATACAGCAATGAAGAGGTGAGAGTCACTC 1620
Db 1561 CCAGAGTCGGAAGAGAAAGCTGAGTTATACAGCAATGAAGAGGTGAGAGTCACTC 1620
Qy 1621 TGAATCACTTGTGTGATGATGAGCCCAACCGCGCAAGGCATGAGAAACAAGGCG 1680
Db 1621 TGAATCACTTGTGTGATGATGAGCCCAACCGCGCAAGGCATGAGAAACAAGGCG 1680
Qy 1681 CTGGGGCACTATGATGTCGCAAGAGGATGGCTGAGAGTTGTGGGCTTGAAGGCTCAA 1740
Db 1681 CTGGGGCACTATGATGTCGCAAGAGGATGGCTGAGAGTTGTGGGCTTGAAGGCTCAA 1740
Qy 1741 TCTGGGGGACCTCAAAATTAATGCTTGGGTGATTCAGATGACCACTAGGCTTGTGT 1800
Db 1741 TCTGGGGGACCTCAAAATTAATGCTTGGGTGATTCAGATGACCACTAGGCTTGTGT 1800
Qy 1801 TGCATGAGTTAAGGATGACAAAGTGAATGAATGAAGTGGGAGAAACAGAAATACCA 1860
Db 1801 TGCATGAGTTAAGGATGACAAAGTGAATGAATGAAGTGGGAGAAACAGAAATACCA 1860
Qy 1861 GCTGTGTCTCAAGGACCACTGAGAGAGAGAAAGAAATGAATGACCACTAGGAGCAC 1920
Db 1861 GCTGTGTCTCAAGGACCACTGAGAGAGAGAAAGAAATGAATGACCACTAGGAGCAC 1920
Qy 1921 ATTGCAAGAACAGTCCCTGGGAGTCTTGTGAGAGCTCAGGAGCTTGTGCTGGCAGAG 1980
Db 1921 ATTGCAAGAACAGTCCCTGGGAGTCTTGTGAGAGCTCAGGAGCTTGTGCTGGCAGAG 1980
Qy 1981 GGAATCTGGCTTACCCAAATTAAGCTCTTGGGATCTGACACATCTAGACCAAGAAATGCA 2040
Db 1981 GGAATCTGGCTTACCCAAATTAAGCTCTTGGGATCTGACACATCTAGACCAAGAAATGCA 2040
Qy 2041 CTGGCAAGAGGTTGAGAGTCTTGGTATTTTCTGATCAATACTGTAAGGCTTTGGG 2100
Db 2041 CTGGCAAGAGGTTGAGAGTCTTGGTATTTTCTGATCAATACTGTAAGGCTTTGGG 2100
Qy 2101 TTCCAAATTTGCTGACAGGCTTTTAACTGAGAAATGGGCTTGAATGAGTATGT 2160
Db 2101 TTCCAAATTTGCTGACAGGCTTTTAACTGAGAAATGGGCTTGAATGAGTATGT 2160
Qy 2161 CTGAGCAAGGCGCTTATCTTCACTGCTTTTCAACCAAGCTGAGGTTGAGAGCTGCT 2220
Db 2161 CTGAGCAAGGCGCTTATCTTCACTGCTTTTCAACCAAGCTGAGGTTGAGAGCTGCT 2220
Qy 2221 TGTTCAGGCTCAAAAATTAAGCTGAGTTTCCAGCAGAGAGGCGCTTATCTGAGCTTCCG 2280
Db 2221 TGTTCAGGCTCAAAAATTAAGCTGAGTTTCCAGCAGAGAGGCGCTTATCTGAGCTTCCG 2280
Qy 2281 TGTCTAGGCTCAATTTCTTCTTCTGTAATAATGACACATGCAACCACTTCCAGTGA 2340
Db 2281 TGTCTAGGCTCAATTTCTTCTTCTGTAATAATGACACATGCAACCACTTCCAGTGA 2340
Qy 2341 CAATGAATATAGACTCAAAACCAATCCCTTGAACCTGCTTGGAGAGGGGCTTGGAGCTAG 2400
Db 2341 CAATGAATATAGACTCAAAACCAATCCCTTGAACCTGCTTGGAGAGGGGCTTGGAGCTAG 2400

Result No.	Score	Query Match	Length	DB	ID	Description
1	3186	100.0	3186	13	US-09-900-448-1	Sequence 1, App11
2	2279.2	71.5	13757	13	US-09-900-448-1	Sequence 3, App11
3	1129.8	35.5	1631	15	US-10-125-237-19	Sequence 19, App11
4	1129.8	35.5	1631	15	US-10-105-831-19	Sequence 19, App11
5	576.8	18.1	1516	13	US-10-312-253-39	Sequence 39, App11
6	576.8	18.1	1516	15	US-10-175-523-84	Sequence 84, App11
7	427	13.2	488	11	US-09-918-995-30828	Sequence 30828, App11
8	427	13.2	488	11	US-09-918-995-8595	Sequence 8595, App11
9	417	13.1	492	11	US-09-918-995-11395	Sequence 11395, App11
10	374.2	11.7	488	11	US-09-918-995-10754	Sequence 10754, App11
11	345.4	10.8	491	11	US-09-918-995-32820	Sequence 32820, App11
12	313.4	9.8	504	10	US-09-880-107-1166	Sequence 1166, App11
13	298.4	9.4	420	10	US-09-960-352-10523	Sequence 10523, App11
14	280	8.8	408	10	US-09-960-352-7040	Sequence 7040, App11
15	278	8.7	384	11	US-09-918-995-7738	Sequence 7738, App11

16	269.4	8.7	472	11	US-09-918-995-32181	Sequence 32181, Ap
17	269.2	8.4	374	10	US-09-960-352-5368	Sequence 5368, Ap
18	266.8	8.4	383	10	US-09-960-352-6886	Sequence 6886, Ap
19	261.4	8.3	437	10	US-09-960-352-380	Sequence 380, App
20	261.8	8.2	405	10	US-09-960-352-445	Sequence 445, App
21	260.2	8.1	425	10	US-09-960-352-3702	Sequence 3702, Ap
22	259	8.2	425	10	US-09-960-352-9378	Sequence 9378, App
23	257.4	8.1	397	10	US-09-960-352-11936	Sequence 11936, A
24	256	8.0	387	10	US-09-960-352-11316	Sequence 11316, A
25	256	8.0	422	10	US-09-960-352-7056	Sequence 7056, App
26	255.2	8.0	418	10	US-09-960-352-13395	Sequence 13395, A
27	255	8.0	373	10	US-09-960-352-2026	Sequence 2026, Ap
28	254.4	8.0	412	10	US-09-960-352-12507	Sequence 12507, App
29	251.4	7.9	435	10	US-09-960-352-13402	Sequence 13402, A
30	251.2	7.9	404	10	US-09-960-352-7514	Sequence 7514, A
31	250.2	7.9	412	10	US-09-960-352-8662	Sequence 8662, Ap
32	250.2	7.9	417	10	US-09-960-352-14176	Sequence 14176, A
33	250.2	7.9	431	10	US-09-960-352-71320	Sequence 71320, App
34	249.8	7.8	399	10	US-09-960-352-457	Sequence 457, App
35	249.4	7.8	407	10	US-09-960-352-136	Sequence 136, App
36	249.4	7.8	410	10	US-09-960-352-6107	Sequence 6107, Ap
37	249.4	7.8	411	10	US-09-960-352-5619	Sequence 5619, App
38	249.2	7.8	366	10	US-09-960-352-44209	Sequence 44209, App
39	248.6	7.8	407	10	US-09-960-352-5920	Sequence 5920, App
40	248.2	7.8	405	10	US-09-960-352-2956	Sequence 2956, App
41	244.8	7.7	407	10	US-09-960-352-13635	Sequence 13635, A
42	244.6	7.7	408	10	US-09-960-352-7025	Sequence 7025, Ap
43	243.6	7.6	392	10	US-09-960-352-911789	Sequence 1789, Ap
44	243.2	7.6	384	10	US-09-960-352-91132	Sequence 9132, Ap
45	239.4	7.5	397	10	US-09-960-352-4643	Sequence 4643, App

ALIGNMENTS

```

RESULT 1
US-09-900-448-1
; Sequence 1, Application US/09900448
; Publication No. US20030220488A1
; GENERAL INFORMATION:
; APPLICANT: CECCARDI, Toni et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1.001272
; CURRENT APPLICATION NUMBER: US/09/900,448
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3186
; TYPE: DNA
; ORGANISM: Human
; US-09-900-448-1

```

		100.0%:	Score 3186:	DB 13:	Length 3186:
		Best Local Similarity	100.0%:	Pred. No. 0:	
		Matches 3186:	Conservative	0:	Mismatches 0:
				Indels 0:	Gaps 0
QY	1 CTCTGAGCTCAGCATGGCTAAGGTACTGGAGACACCCTGTTCACCTGGGTTTGTGAACC	60			
Db	1 CTCCTGAGGTCAGCATGGCTTAGGGTACTGGAGACACCCTGTTCACCTGGGTTTGTGAAGCC	60			
QY	61 TAATGCTGCTCTTGAGCCATTGGCACCCCTCTTCTCCGCACTAATGAGGCCATGGGAATGTTG	120			
Db	61 TAATGCTGCTCTTGAGCCATTGGCACCCCTCTTCTCCGCACTAATGAGGCCATGGGAATGTTG	120			
QY	121 CTGAAGGCAGACCAAAGCGACAACCCAGACGTCATGAACCGTGTCCAGATGGCTGAGAAGCT	180			
Db	121 CTGAAGGCAGACCAAAGCGACAACCCAGACGTCATGAACCGTGTCCAGATGGCTGAGAAGCT	180			
QY	181 TTGATGCTACACCCCTGGATGACAAATGGACCAATGCTGTGTTTTTAAAGGCGAGTTGTGT	240			

QY	3007	TGTTTACGTCAGTTTAGGCGTCTATAGAGAAATATCTAGAGGGTAACTATCAGCAA	3066
Db	37631	TGCTTAGTCGTGTTTGCTGCTATACACAATACCTAGACTAACTACTTAT--AAC	37575
QY	3067	TAGAAATTTATTTGTTACAAATTCGTGGAGGCTGGAAAATCCAAATCAAGGCTCCAGCAG	3126
Db	37574	TGGAATCTCTTCTTCACAGTTCTTGCAAGCCGGGAAGTCCAAATCAAGCTCTGGCAG	37515
QY	3127	TTTCAG-TGTCGTGCTAGTCTGTGT-TCTGCTTCGAAGATGGCACTTTTTCGTGTTCT	3184
Db	37514	TTTACGTGTCGTGTGGAGGAGCAGCTCTTCATCCAAAGATGGACCTTGTGTGCATCTT	37455
QY	3185	C	
Db	37454	C	37454

RESULT 15
US-09-877-177A-10/c
: Sequence 10, Application us/09877177A
: Patent No. 6582919
: GENERAL INFORMATION:
: APPLICANT: K. Danenberg
: TITLE OF INVENTION: Method of determining Epidermal Growth
: TITLE OF INVENTION: Factor Receptor and HBR2-New Gene Expression
: TITLE OF INVENTION: and Correlation of Levels Therof With Survival
: FILE REFERENCE: 11220/120
: CURRENT APPLICATION NUMBER: US/09/877,177A
: CURRENT FILING DATE: 2001-06-11

```

RESULT 15
US-09-877-177A-10/C
; Sequence 10; Application US/09877177A
; Patent No. 6582819
; GENERAL INFORMATION
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-10

Query Match          2.1%; Score 68.4; DB 4; Length 197496;
Best Local Similarity 67.6%; Pred. No. 3.6e-10;
Matches 96; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY      3004 GTTGTTCAGTCAGTTAGGCTGCTATAAGAGATATCTTAGAGGTGGTATCTATCAG 3063
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3500 GTATGCTTCATCTTTTGGGCTTCGTAAACAATAATACATTAATATGATTTACTGTAA 34411

```

```

Query Match      2.1%; Score 68.4; DB 4; Length 197496;
Best Local Similarity 67.6%; Pred. No. 3.6e-10;
Matches 96; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY      3004  GTTGTTTCAGTCAGTTAGGCTGCTCTATAAGAGAAATCTTAGAGTGGGTATCTATCAG 3063
DB      3500  GTATGCTTATCTTTTGGGCTCTCTAACAATAATTAATATTAATTACTGTATA 3441

QY      3064  CAATAGCAATTTATTTGTTCACAANTCTGAGGCTGGAATAATCCAAGATCCAAGCTCCAGC 3123
DB      3440  CAACAGAAATTTATTTCTCCATPAGTTCTGAGAGCTGGGAATTTCAAGATCCAAGTATTTGC 3381

QY      3124  AGGTTCAGTGCTGCTGAGTGC 3145
DB      3380  AGATTTGATGTCTATATNAGGCG 3359

Search completed: December 16, 2003, 05:46:23
Job time : 141.986 secs

```

Search completed: December 16, 2003, 05:46:23
job time : 141.986 secs

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 69.8; DB 4; Length 762;

Best Local Similarity 71.3%; Pred. No. 2,4e-12; Matches 92; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 3058 TATCAGCAATAGCAATTTATTTGTTCAATTTGAGGCTGGAAATCCAGATCAAGGC 3117
DB 756 TATTAACAGCAAAATATTTATTTCTTACAGTTCTAGAGCTGGGAAGTCCAAGATCAAGGC 697
QY 3118 TCCAGCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3177
DB 696 ACCAGCAGATTCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637
QY 3178 GTGTTCTCA 3186
DB 636 GTGATCACA 628

RESULT 13
US-09-205-258-197/c

;; Sequence 197, Application US/09205258
;; Patent No. 6525174
;; GENERAL INFORMATION:
;; APPLICANT: Young et al.
;; TITLE OF INVENTION: 207 Human Secreted Proteins
;; FILE REFERENCE: P2007P1
;; CURRENT FILING DATE: 1998-12-04
;; EARLIER APPLICATION NUMBER: PCT/US98/11422
;; EARLIER FILING DATE: 1998-06-04
;; EARLIER APPLICATION NUMBER: 60/048,885
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/049,375
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,881
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,880
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,896
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/049,020
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,876
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,895
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,884
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,894
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,971
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,964
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,882
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,899
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,893
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,900
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,901
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,892
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,915
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/049,019
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,970
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,972
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,916
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/049,373
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,875
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/049,374
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,917
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,949
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,974
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,883
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,897
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,898

Patent No. 6478825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C3
CURRENT APPLICATION NUMBER: US/09/996,243
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429

Db 900 ACAGAGATGATGCTCAGAGTCTGGAGGCTGGAGATTCAGATCAATGACCCGAG 959
QY 3126 GTTCAGTCTGCTGAGATGCTGTTCT 3152
Db 960 TTTCGGTGTGTGGAGAGCTTTGCT 986

RESULT 10

US-08-611-587-1/c
Sequence 1, Application US/08611587
Patent No. 6150091
GENERAL INFORMATION:
APPLICANT: PANDOLFO, MASSIMO
APPLICANT: MONTERMINI, LAURA
APPLICANT: MOLTO, MARIA D.
APPLICANT: Koenig, Michael
APPLICANT: Campuzano, Victoria
TITLE OF INVENTION: Direct Diagnosis of Friedreich Ataxia
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P. Patent Dept.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.
ZIP: 77010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,587
FILING DATE: 03-MAR-1996
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Brashears-Macatee, Sarah J.
REGISTRATION NUMBER: 38,087
REFERENCE/DOCKET NUMBER: D-5901
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5620
TELEFAX: 713-651-5246
TELEX: 76-2829
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 9q13
UNITS: bp
US-08-611-587-1

Query Match 2.5%; Score 78.4; DB 3; Length 8353;
Best Local Similarity 66.8%; Pred. No. 1,9e-14;
Matches 12; Conservative 0; Mismatches 61; Indels 2; Gaps 1;

QY 2996 CAGAAATGTTGTTGTTGAGTCAAGTCTGCTATAAGAGAAATATCTTAGAGTGGTAA 3055
Db 8244 CAAATATTTGTTATTTGTTGTTAGTCTGGAGTGTATTAACAAATATACAGCACTGGGTGG 8185
QY 3056 TCATATCGCAATGAGATTTATTTGTTTCAATTTCTGAGAGCTGGAATATCCAGATCAAG 3115
Db 8184 CTTATTAACCAATGAGAAATGTTGTTCTTAGAGTCTCGAGAGGTGGGAATCAAGATCAAT 8125

QY 3116 GCTCCAGAGATTCAGTGTCTGCTGAGTCTTGT--CTGCTTGAAGATGGACCTTTT 3173
Db 8124 GCACCAACAGCTTTGGTGTCTGGTGAAGGACAGTTTCTGTGTTCAATAGATGACCTTGT 8065
QY 3174 TGCTGTGTTT 3183
Db 8064 CACTGGGCTC 8055

RESULT 11

US-09-009-913-1/c
Sequence 1, Application US/09009913
Patent No. 6087485
GENERAL INFORMATION:
APPLICANT: AYS Pharmaceuticals, Inc.
TITLE OF INVENTION: Asthma Related Genes
NUMBER OF SEQUENCES: 339
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,913
FILING DATE: 21-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-4P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3231
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 72928 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-009-913-1

Query Match 2.4%; Score 77.6; DB 3; Length 72928;
Best Local Similarity 72.2%; Pred. No. 1.6e-13;
Matches 130; Conservative 0; Mismatches 44; Indels 6; Gaps 2;

QY 3009 TTTCAGTCAAGTCTGCTGCTATTAAGAGATATCTTAGAGTGGTATCTATACGAATA 3068
Db 36263 TCTTAGTCAATTTACCTGCAACAAAGATGCTAGACTAGG---CTACAGCAACG 36208
QY 3069 GGAATTTATTTGTTCAATTTCTGAGAGCTGGAATATCCAGATCAAGGCTCCAGAGTT 3128
Db 36207 GAAATTTATTTATTTCAATTTCTAGAGCTGGAGATGCCAAGACCAAGCAGATTT 36148
QY 3129 CAGTGTCTGTAGTG--CTGTCTGCTTGAAGATGGACCTTTTGTGTGTTCTCA 3186
Db 36147 CAGCGTGTGTAGGGCCCACTTCTTAGTTCATAGATGCATCTTCTCACCGTGAACCTCA 36088

RESULT 12
US-09-996-243-344/c
Sequence 344, Application US/09996243

EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentn Ver. 2.0
SEQ ID NO 4
LENGTH: 50000
TYPE: DNA
ORGANISM: Homo sapiens
US-09-146-053-4

Query Match 2.5%; Score 80.4; DB 4; Length 50000;
Best Local Similarity 68.7%; Pred. No. 1.5e-14;
Matches 125; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY 3005 TTGTTTCACTAGTTAGGCTGCTATAGAGATATCTTAGTGGTATCTATCAGC 3064
DB 37896 TTATATCTTATCTCATTTGGGCTGCTATATACAAAGCCTGCAGACTGTTGGCTTATTAAC 37955
QY 3065 AATAGAAATTATTTATTTTCAATTTCTGAGGCTGGAATAATCCAAAGATCAAGGCTCCAGCA 3124
DB 37956 AGTAGACATTATTTCTTACAGTTCTGAGGCTGGGAAGTCGAAGATCAAGGACCTAGCA 38015
QY 3125 GATTGAGTCTGCTGAGTCTGTTCTTCTGGAAGATGACACTTTTCTGCTGTTCT 3184
DB 38016 GATTGAGTCTGCAACAGGCGCCAGT-TCTTCTGATAGTGCACCTCTAGCTGTATCTT 38074
QY 3185 CA 3186
DB 38075 CA 38076

RESULT 8

US-09-671-317-212
Sequence 212, Application US/09671317
Patent No. 6528260
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIOMOLECULAR MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
FILE REFERENCE: 62, US3, CIP
CURRENT APPLICATION NUMBER: US/09/671,317
CURRENT FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 212
LENGTH: 1001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 501
OTHER INFORMATION: 12-500-220 : polymorphic base A or G
NAME/KEY: misc binding
LOCATION: 481..500
OTHER INFORMATION: 12-500-220. misc2, potential complement
NAME/KEY: primer bind
LOCATION: 283..303
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer bind
LOCATION: 711..731
OTHER INFORMATION: downstream amplification primer, complement

NAME/KEY: misc binding
LOCATION: 489..513
OTHER INFORMATION: 12-500-220 potential probe
US-09-671-317-212

Query Match 2.5%; Score 79; DB 4; Length 1001;
Best Local Similarity 75.5%; Pred. No. 2.7e-15;
Matches 111; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 3007 TGTTCAGTCAGTTTGGCTGCTATAGA-GAATATCTTAGAGTGGTATATCATCAGCA 3065
DB 837 TGTTCAGTTTGTGGCTGCTGTAACAAATATCTTACAGTGGTGGCTTATTAACA 896
QY 3066 ATAGAAATTATTTATTTTCAATTTCTGAGGCTGGAATAATCCAAAGATCAAGGCTCCAGCAG 3125
DB 897 ACGGAATGATTTGCTCAGAGTTCTGAGGCTGGGAAGTTACGATCAAGTACCGGCGAG 956
QY 3126 GTTCACTGCTGCTGAGTCTGTTCTTCT 3152
DB 957 TTTCGAGTGTGTGAGAGCTTTTCT 983

RESULT 9

US-09-671-317-458
Sequence 458, Application US/09671317
Patent No. 6528260
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIOMOLECULAR MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
FILE REFERENCE: 62, US3, CIP
CURRENT APPLICATION NUMBER: US/09/671,317
CURRENT FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 458
LENGTH: 1001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 501
OTHER INFORMATION: 12-500-217 : insertion CAATA
NAME/KEY: misc binding
LOCATION: 481..500
OTHER INFORMATION: 12-500-217. misc1, potential
NAME/KEY: primer bind
LOCATION: 286..306
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer bind
LOCATION: 714..734
OTHER INFORMATION: downstream amplification primer, complement
US-09-671-317-458

Query Match 2.5%; Score 79; DB 4; Length 1001;
Best Local Similarity 75.5%; Pred. No. 2.7e-15;
Matches 111; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 3007 TGTTCAGTCAGTTTGGCTGCTATAGA-GAATATCTTAGAGTGGTATATCATCAGCA 3065
DB 840 TGTTCAGTTTGTGGCTGCTGTAACAAATATCTTACAGTGGTGGCTTATTAACA 899
QY 3066 AATAGAAATTATTTTCAATTTCTGAGGCTGGAATAATCCAAAGATCAAGGCTCCAGCAG 3125

LOCATION: 68697
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68718
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68733
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68739
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 69785
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79134
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79198
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 86336
OTHER INFORMATION: unknown
OTHER INFORMATION: unknown
US-09-791-211-3

Query Match 2.7%; Score 86; DB 4; Length 87543;
Best Local Similarity 73.3%; Pred. No. 3.1e-16;
Matches 151; Conservative 0; Mismatches 50; Indels 5; Gaps 3;

QY 2983 TCAGGGAATCCAGAAATGTTGTTTCAGTCAGTTAGGCTGCT--ATAAGGAAT 3039
DB 66969 TACTAGATGTGCGACAGATGATTTGCTTAGTCTGTTGCTGTTTATACAAAT 68910
QY 3040 ATCTTAGAGTGGTAACTATTCAGCATAGAAATTTATGTTTCAATTCGAGGCTGG 3099
DB 68909 GCCACAGACTGGGCAATTTCAACAACAGATGTTATTTCTATAGTTCTGGAGGCTGG 68850
QY 3100 AAAATCCAGATCAGGCTCCAGAGGTTCACTGCTGCTGAGTCTTGT-TCTGCTTCG 3158
DB 68849 -AGGTCAGAGATCAAGGCGCCAGTAGGTTCACTGCTGAGGAGGCTGCTCTGCTTCC 68791
QY 3159 AAGATGGACCTTTTGGCTGCTGTTCT 3184
DB 68790 AAAATGAGGCTGTGCTGTGTCT 68765

RESULT 5
US-09-804-471A-3

Sequence 3, Application US/09804471A
Patent No. 6479269
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL001164
CURRENT APPLICATION NUMBER: US/09/804,471A
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 174493
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(174493)
OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3

Query Match 2.6%; Score 82.6; DB 4; Length 174493;
Best Local Similarity 72.2%; Pred. No. 6.8e-15;

Matches 135; Conservative 0; Mismatches 49; Indels 3; Gaps 2;
QY 3000 AATGTTTGTTCAGTCAGTTAGGCTGCTATTAAGAAATATCTTAGAGTGGTATCTA 3059
DB 14097 AAACGTGATCTGAGCTGTGTTGGGCTGCAATATAG-ATACCTTAGACTTGGCAATTTA 14155
QY 3060 TCAGCAATAGAAATTAATTTGTTACAAATCTGAGGCTGGAATAATCAAGATCAAGCTC 3119
DB 14156 TAAACATAGAAATTCATTGCTGACAGTTGTGAAGCTGGAAAGTCCAAAGATCAAGCCG 14215
QY 3120 CAGCAGTTCAGTGTCTGCTGAGTGGCTTGTCTTTCGAAGATGACCTTTTGGCTGT 3179
DB 14216 CAGCGAATCTGATCTGTGATAGGCT--CCCTGCTTCAAAATGCGGCTTCTTCTGCG 14273
QY 3180 GTTCTCA 3186
DB 14274 ATCTTCA 14280

RESULT 6
US-09-873-404-3

Sequence 3, Application US/09873404
Patent No. 6500656
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THERBOF
FILE REFERENCE: CL001212-CIP
CURRENT APPLICATION NUMBER: US/09/873,404
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 63588
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(63588)
OTHER INFORMATION: n = A,T,C or G
US-09-873-404-3

Query Match 2.5%; Score 80.8; DB 4; Length 63588;
Best Local Similarity 71.6%; Pred. No. 1.3e-14;
Matches 106; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2999 AATGTTTGTTCAGTCAGTTAGGCTGCTATTAAGAAATATCTTAGAGTGGTATCT 3058
DB 15658 AATCATCTGTTTACTATTCACAGACTATTAACAATAATCAATTAAGTGGCTT 15717
QY 3059 ATCAGCAATAGAAATTAATTTGTTACAAATCTGAGGCTGGAATAATCAAGATCAAGCT 3118
DB 15718 ATTAACAGCAAAATTAATTTGTCGAGTTCGAGAGCTGGAAAGTCAAGATCAAGCTG 15777
QY 3119 CAGCAGTTCAGTGTCTGCTGAGTCT 3146
DB 15778 CAGCAGATTCAGTGTCTGAGGCTT 15805

RESULT 7
US-09-146-053-4

Sequence 4, Application US/09146053A
Patent No. 6399349
GENERAL INFORMATION:
APPLICANT: Ryan, James W.
APPLICANT: Sprinkle, Terry Joe Curtis
APPLICANT: Venema, Richard C.
TITLE OF INVENTION: Human Aminopeptidase P Gene
FILE REFERENCE: MCG103
CURRENT APPLICATION NUMBER: US/09/146,053A
CURRENT FILING DATE: 1998-09-02
EARLIER APPLICATION NUMBER: 60/057,854

NAME/KEY:	unsure
LOCATION:	31206
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	31592
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	33095
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	33160
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	34066
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	34072
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	36816
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	39020
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	42154
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	42459
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	46808
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	46823
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	47291
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	52786
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	52787
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	53384
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	54684
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	59215
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	59235
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	59242
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	63280
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	66614
OTHER INFORMATION:	unknown
NAME/KEY:	unsure
LOCATION:	68660
OTHER INFORMATION:	unknown
NAME/KEY:	unsure

OTHER INFORMATION: 99-15668-139.m162, complement
FEATURE:
NAME/KEY: primer bind
LOCATION: 1363..1380
OTHER INFORMATION: upstream amplification primer
FEATURE:
NAME/KEY: primer bind
LOCATION: 1801..1821
OTHER INFORMATION: downstream amplification primer, complement
FEATURE:
NAME/KEY: misc binding
LOCATION: 1489..1513
OTHER INFORMATION: 99-15668-139 probe
US-09-539-333D-177

Query Match 2.8%; Score 88; DB 4; Length 3001;
Best Local Similarity 71.7%; Pred. No. 6.2e-18;
Matches 129; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

QY 3008 GTTTCAGTCAGTTAGGCTGCTATAGAGAAATCTTAGTGAGTGGTAACTATCAGCAAT 3067
DB 2055 GTTCTAGTCAGTTTGTGCTGCTGTAACAGAAATCCACAGCTGGCTATTAAATGAAGAG 2114
QY 3068 AGGAATTTATTTGTTCAATTTGAGAGCTGGAATAATCCAAAGATCAAGCTCCAGCAGGT 3127
DB 2115 AAGAAATTCATTTCTCATAGTTGTGAGAGCTAGAAAGTCCAGATGAAGGCGCCAGCAAT 2174
QY 3128 TCAGTGTCTGCTGAGTGTCT-TGTTCTGCTTCGAAGATGGACCTTTTGTGTTCTCA 3186
DB 2175 TTGTTGTGTGTGAGGGGTGCGCTGCTTCCAAGATGAGCTTTGATGCTGCAATTTTCA 2234

RESULT 2

US-08-781-891-79/c
Sequence 79, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-Bn
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6090620cendurg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 622-6031
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 87350 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
US-08-781-891-79

Query Match 2.7%; Score 86; DB 3; Length 87350;
Best Local Similarity 73.3%; Pred. No. 3.1e-16;
Matches 151; Conservative 0; Mismatches 50; Indels 5; Gaps 3;

QY 2983 TCCAGGAATCCCGAAGATGTTGTTCACTGAGTCTGCT--ATPAGCAAT 3039
DB 68776 TACTAGTATGCGCAAGACTGATGTCTTAGTCTGTTTCTGTTTCATPACAAAT 68717
QY 3040 ACTTAGTGGGTATATATATAGCAATAGATTTATTTTCAATTTCTGAGGCTG 3099
DB 68716 GCCACAGACTGGGCAATTTCAACACAGTATTTATTTTCATAGTCTGAGGCTG 68657
QY 3100 AAAATCAAGATCAAGCTCCAGCAGTTCAGTGTCTGAGTCTGT-TCTGCTTCG 3158
DB 68656 -AGTCCAGATCAAGGCGCCAGTAGTTCAGTGTCTGAGGCTGCTCTCTCC 68598
QY 3159 AAGATGCACTTTTGTCTGTCTTCT 3184
DB 68597 AAAATGAGCTTGTGCTGTCTTCT 68572

RESULT 3

US-09-618-166-79/c
Sequence 79, Application US/09618166
Patent No. 6583112

GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-Bn
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419c1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 622-6031
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 87350 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-618-166-79

Query Match 2.7%; Score 86; DB 4; Length 87350;
Best Local Similarity 73.3%; Pred. No. 3.1e-16;
Matches 151; Conservative 0; Mismatches 50; Indels 5; Gaps 3;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 15, 2003, 11:21:38 ; Search time 134.986 Seconds
(without alignments)
10417.750 Million cell updates/sec

Title: US-09-900-448-1

Perfect score: 3186
Sequence: 1 cctcgagctcagcatgct.....accttttgcgtgtctcca 3186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :
1: Issued Patents NA:*
2: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
4: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
5: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	2.8	3001	US-09-539-333D-177	Sequence 177, App
2	86	2.7	87350	US-08-781-891-79	Sequence 79, Appl
3	86	2.7	87350	US-09-618-166-79	Sequence 79, Appl
4	86	2.7	87543	US-09-791-211-3	Sequence 3, Appl1
5	82.6	2.6	174493	US-09-804-471A-3	Sequence 3, Appl1
6	80.8	2.5	63588	US-09-873-404-3	Sequence 3, Appl1
7	80.4	2.5	50000	US-09-146-053-4	Sequence 4, Appl1
8	79	2.5	1001	US-09-671-317-212	Sequence 212, App
9	79	2.5	1001	US-08-611-587-1	Sequence 458, App
10	78.4	2.5	8353	US-09-009-913-1	Sequence 1, Appl1
11	77.6	2.4	72928	US-09-096-243-344	Sequence 344, App
12	69.8	2.2	762	US-09-396-243-344	Sequence 197, App
13	69.2	2.2	1282	US-09-205-258-197	Sequence 10, Appl
14	69	2.2	50000	US-09-146-053-3	Sequence 145, App
15	68.4	2.1	197496	US-09-877-177A-10	Sequence 145, App
16	68.2	2.1	309	US-09-322-575-145	Sequence 145, App
17	68.2	2.1	309	US-09-389-681-145	Sequence 145, App
18	68.2	2.1	309	US-09-620-405B-145	Sequence 145, App
19	68.2	2.1	309	US-09-339-338-145	Sequence 145, App
20	68.2	2.1	309	US-09-433-826B-145	Sequence 145, App
21	68.2	2.1	309	US-09-604-287A-145	Sequence 145, App
22	67.8	2.1	50000	US-09-146-053-4	Sequence 4, Appl1
23	66.8	2.1	41684	US-09-536-059-1	Sequence 1, Appl1
24	66.8	2.1	1001	US-09-641-638-459	Sequence 459, App
25	65.8	2.1	7218	US-08-232-463-14	Sequence 14, Appl
26	63.2	2.0	246240	US-08-724-394A-20	Sequence 20, Appl
27	63.2	2.0	246240	US-08-724-394A-21	Sequence 21, Appl

C 28	63.2	2.0	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 29	63	2.0	3001	4	US-09-539-333D-192	Sequence 192, App
C 30	62.8	2.0	44453	4	US-09-146-053-5	Sequence 5, Appl
C 31	62	1.9	319608	4	US-09-539-333D-1	Sequence 1, Appl1
C 32	62	1.9	319608	4	US-09-679-403-1	Sequence 1, Appl1
C 33	61.8	1.9	152331	3	US-09-128-155-16	Sequence 16, Appl
C 34	61.8	1.9	176373	3	US-09-128-155-17	Sequence 17, Appl
C 35	61.4	1.9	8355	3	US-08-406-030A-23	Sequence 23, Appl
C 36	60.8	1.9	168575	4	US-09-426-290-1	Sequence 138, App
C 37	60.4	1.9	3001	4	US-09-539-333D-138	Sequence 3, Appl1
C 38	59.4	1.9	202001	4	US-09-734-674-3	Sequence 3, Appl1
C 39	59	1.9	80246	3	US-09-078-294-4	Sequence 4, Appl1
C 40	59	1.9	80595	3	US-09-078-294-3	Sequence 3, Appl1
C 41	57	1.8	49312	4	US-09-671-317-485	Sequence 485, App
C 42	55	1.7	116592	4	US-09-818-512-3	Sequence 3, Appl1
C 43	53.6	1.7	49136	3	US-09-422-865-1	Sequence 1, Appl1
C 44	52	1.6	43069	4	US-09-292-542A-1	Sequence 1, Appl1
C 45	51	1.6	1815	4	US-09-328-475C-339	Sequence 339, App

ALIGNMENTS

RESULT 1
US-09-539-333D-177
Sequence 177, Application US/09539333D
Patent No. 6476208
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Bihain, Bernard
APPLICANT: Essioux, Laurent
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLIC MARKERS
FILE REFERENCE: GENSET.047AUS
CURRENT APPLICATION NUMBER: US/09/539,333D
CURRENT FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US 60/126,903
PRIOR FILING DATE: 1999-03-30
PRIOR APPLICATION NUMBER: US 60/131,971
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: US 60/132,065
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: US 60/143,928
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: US 60/145,915
PRIOR FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: US 60/146,453
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: US 60/146,452
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: US 60/162,288
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: US 09/416,384
PRIOR FILING DATE: 1999-10-12
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patent.pm
SEQ ID NO 177
LENGTH: 3001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 1501
OTHER INFORMATION: 99-15668-139 : polymorphic base C or T
FEATURE:
NAME/KEY: misc binding
LOCATION: 1482..1500
OTHER INFORMATION: 99-15668-139.misl
FEATURE:
NAME/KEY: misc binding
LOCATION: 1502..1521

XX Claim 2; SEQ ID No 445; 245bp; English.

PS The invention relates to a purified nucleic acid molecule associated with

XX laccation or muscle and fat deposition (designated LMPD), derived

CC from cattle, and the LMPD nucleic acid can specifically hybridise to a

CC second nucleic acid molecule comprising any of 15112 nucleotide

CC sequences, appearing as ABX34836-ABX49947, or complements of them.

CC Also included are: (1) a transformed cell having a nucleic acid

CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-

CC translated sequence that functions in the cell to cause termination of

CC transcription and addition of polyadenylated ribonucleotides to a 3' end

CC of the mRNA molecule; and (2) determining a level or pattern of a

CC molecule in a bovine cell or tissue comprising: (a) incubating a marker

CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its

CC complement or fragment) with a complementary nucleic acid molecule

CC obtained from the bovine cell or tissue, where hybridisation between the

CC marker nucleic acid and the complementary nucleic acid permits the

CC detection of the molecule; and (b) detecting the level or pattern of the

CC complementary nucleic acid, where the detection of the complementary

CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMPD nucleic acid is used for determining a level or pattern

CC of a molecule in a bovine cell or tissue. It is useful for genome

CC mapping, gene identification and analysis, cattle breeding, preparation

CC of constructs for use in cattle gene expression, or for genetically

CC improving cattle. The present sequence is one of the 15112 bovine

CC LMPD EST (expressed sequence tag) nucleic acids.

CC Note: The present sequence was not shown in the specification but

CC was obtained in electronic format from the USPTO web site:

CC seqdata.uspto.gov/sequence.html?docid=20020137139.

XX Sequence 405 BP; 93 A; 91 C; 130 G; 91 T; 0 other;

SQ

Query Match 8.2%; Score 261.8; DB 25; Length 405;

Best Local Similarity 81.3%; Pred. No. 2.9e-73;

Matches 331; Conservative 0; Mismatches 67; Indels 9; Gaps 2;

QY 75 GCCATGGCCACCCCTCTTCCTCCGACGAGCCCATGAGGAAATGTCGTAAGGCGAGA--132

DB 2 GCTTAAGCCCACTCTGCTGCTAGAGCTCTGAGCTTGGGACGAGTGAAGGTGGGAAT 61

QY 133 ----CCAAGCAGACCCAGACGTAAGTGAACGCTGCTCAAGATGCTGAGACTTGTGATGCT 188

DB 62 GTGGCCCAAGCCGATCCAGAAAGTGAAGGAAAGCTGCTGGAGCTTGTGATGCT 121

QY 189 ACCACCTGATGACATGGAACATGCTCTTTTAAAGGGAGTTGTGTGAAGACT 248

DB 122 ACCACCTGATGAGATGGAACATGCTCTTTTAAAGGGAGTTGTGTGAAGACT 181

QY 249 CACCAATGGGACCGGAGTTAATCTCAGAGATGGAAGAAATTTCCCGCCCTGTGAT 308

DB 182 CATGCTGGGCGCCGAGTTGATCTCAGAGATGGAAGAAATTTCCCGCCCTGTGAT 241

QY 309 GCTGATTCCTGCTGAGATGCAACAAGTCTTCTGATCAAGGGGAGCAAAAGTCTGGGTA 368

DB 242 GCTGATTCCTGCTGAGATGCAACAAGTCTTCTGATCAAGGGGAGCAAAATCTGGGTA 301

QY 369 TACCTCTCTGAAAGAGAAAGATATCCCAAGTTGCTCCAAGATGAAATTTCTGGA 428

DB 302 TATCCCTCTG--AGAAAGGGGAGGAATATCCAAAGTTGCTCCAAGATGAAATTTCTGGA 358

QY 429 ATCCCATCCCACTGATGAGCTGTGGAATGTCACCGTGGAGAAATG 475

DB 359 ATCCCATCCCACTGATGAGCTGTGGAATGTCACCGTGGAGAAATG 405

RESULT 14

ABX38537

ID ABX38537 standard; cDNA; 425 BP.

XX ABX38537;

XX 20-FEB-2003 (first entry)

XX Bovine EST associated with laccation/muscle/fat deposition #3702.

DB Bovine; 89; EST; expressed sequence tag; laccation; LMPD;

XX muscle deposition; fat deposition; genome mapping; gene identification;

XX gene analysis; cattle breeding.

XX Bos Taurus.

OS US2002137139-A1.

PN 26-SEP-2002.

XX 24-SEP-2001; 2001US-0960352.

PF 12-JUN-1999; 99US-115707P.

PR 11-JAN-2000; 2000US-0480902.

XX (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

PI WPI; 2003-110599/10.

DR

XX New nucleic acid associated with laccation, and muscle and fat

PT deposition, useful for genome mapping, gene identification and

PT analysis, cattle breeding, or for genetically improving cattle

XX

PS Claim 2; SEQ ID No 3702; 245bp; English.

XX

XX The invention relates to a purified nucleic acid molecule associated with

CC laccation or muscle and fat deposition (designated LMPD), derived

CC from cattle, and the LMPD nucleic acid can specifically hybridise to a

CC second nucleic acid molecule comprising any of 15112 nucleotide

CC sequences, appearing as ABX34836-ABX49947, or complements of them.

CC Also included are: (1) a transformed cell having a nucleic acid

CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-

CC translated sequence that functions in the cell to cause termination of

CC transcription and addition of polyadenylated ribonucleotides to a 3' end

CC of the mRNA molecule; and (2) determining a level or pattern of a

CC molecule in a bovine cell or tissue comprising: (a) incubating a marker

CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its

CC complement or fragment) with a complementary nucleic acid molecule

CC obtained from the bovine cell or tissue, where hybridisation between the

CC marker nucleic acid and the complementary nucleic acid permits the

CC detection of the molecule; and (b) detecting the level or pattern of the

CC complementary nucleic acid, where the detection of the complementary

CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMPD nucleic acid is used for determining a level or pattern

CC of a molecule in a bovine cell or tissue. It is useful for genome

CC mapping, gene identification and analysis, cattle breeding, preparation

CC of constructs for use in cattle gene expression, or for genetically

CC improving cattle. The present sequence is one of the 15112 bovine

CC LMPD EST (expressed sequence tag) nucleic acids.

CC Note: The present sequence was not shown in the specification but

CC was obtained in electronic format from the USPTO web site:

CC seqdata.uspto.gov/sequence.html?docid=20020137139.

XX Sequence 425 BP; 88 A; 95 C; 138 G; 103 T; 1 other;

SQ

Query Match 8.2%; Score 260.2; DB 25; Length 425;

Best Local Similarity 80.4%; Pred. No. 9.7e-73;

Matches 344; Conservative 0; Mismatches 74; Indels 10; Gaps 3;

QY 36 CCCGTGACATGGGCTTGGAGCCCTAAGCTGCTGCGCATGCCACCCCTCTCTCT 95

DB 2 CCAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61

QY 96 CCGACTAGTCCCATGAGGAAATGCTGTAAGGCGAGA-----CCAAGCAGACCGAGAC 149

Matches 308; Conservative 0; Mismatches 53; Indels 3; Gaps 1;

QY 145 CAGACGTGACTGAACGCTGCTCAGATGCTGAGCTTTGATGTACCACTGGATGACA 204
 1 CAGAGTGAAGGAAACGCTGCTCGATGCTGGGCTTTGATGTACCACTGGATGAGC 60
 QY 205 ATGGAACATGCTGCTTTTAAAGGGGAGTTTGTGTGGAAGTCACAAAATGGACCGG 264
 61 ATGGGAACATGCTGCTTTTAAAGGGGAGTTTGTGTGGAAGTCACAAAATGGACCGG 120
 QY 265 AGTTAATCTCAGAGAGATGAGAAATTTCCCAAGCCCTGAGATGCTGCAATCCGTAAG 324
 121 AGTTGATCTCAGAGAGATGAGAAATTTCCCAAGCCCTGAGATGCTGCAATCCGTAAG 180
 QY 325 GTCAACAACAGTCTTTCTGATCAAGGGGAGCAAAAGTCTGGATATCTCTGTAAGA 384
 181 ATGTAAACAGTCTTTCTGATCAAGGGGAGCAAAAGTCTGGATATCTCTGTAAGA 237
 QY 385 AGGAGAAAGATACCCAAAGTTGCTCAAGATGAATTTCTGCAATCCATCCCACTGG 444
 238 AGGGGAGGAAATTCCAAAGTTGCTCAAGATGAATTTCTGCAATCCATCCCACTGG 297
 QY 445 ATGCAAGTGTGAGATGTCACCGTGGAGATGTCAAGCTGAAGGCTCTCTCTTCCAG 504
 298 ATGCAAGTGTGAGATGTCACCGTGGAGATGTCTCAATGAGGATGTCCTCTTCCAG 357
 QY 505 GCCA 508
 358 GCCA 361
 Db

RESULT 12
 AAC00310
 ID AAC00310 standard; cDNA; 279 BP.

XX AAC00310;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 308.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX BP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000BP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Maline Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG00304.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 308; 71bp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.

XX Sequence 279 BP; 56 A; 71 C; 87 G; 64 T; 1 other;

Query Match 8.3%; Score 263.6; DB 21; Length 279;
 Best Local Similarity 99.6%; Pred. No. 5, 9e-74;
 Matches 263; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGTGAGCTCAGCATGAGCTAGGCTAGCTGAGACACCCGTTGCACTGGGTTTGAGCC 60
 16 CTGTGAGCTCAGCATGAGCTAGGCTAGCTGAGACACCCGTTGCACTGGGTTTGAGCC 75
 QY 61 TATGCTGCTCTGAGCATTTGCCACCTCTTCTCCGACTAGTGGCCATGGGAATGTTG 120
 76 TATGCTGCTCTGAGCATTTGCCACCTCTTCTCCGACTAGTGGCCATGGGAATGTTG 135
 Db 121 CTGAAGCCGAGACCAAGCCAGACCTGAGCTGAAAGCTGCTGAGAGCT 180
 136 CTGAAGCCGAGACCAAGCCAGACCTGAGCTGAAAGCTGCTGAGAGCT 195
 QY 181 TTGATGCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 196 TTGATGCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
 QY 241 GGAAGATCACAATGGACCGGG 264
 256 GGAAGATCACAATGGACCGGG 279
 Db

RESULT 13
 ABX35280
 ID ABX35280 standard; cDNA; 405 BP.

XX ABX35280;

XX 20-FEB-2003 (first entry)

XX Bovine EST associated with lactation/muscle/fat deposition #445.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMD;

XX muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137139-A1.

XX 26-SEP-2002.

XX 24-SEP-2001; 2001US-0960352.

XX 12-JAN-1999; 99US-115707P.

XX 11-JAN-2000; 2000US-0480902.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;
 PT WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle -

XX Byatt JC, Mathalagan N, Tao N, Warren WC;
 PI MPI; 2003-110599/10.
 XX
 DR New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX
 PS Claim 2; SEQ ID No 6846; 245bp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are: (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.
 CC
 XX
 SQ Sequence 383 BP; 85 A; 108 C; 94 G; 96 T; 0 other;
 Query Match 8.4%; Score 266.8; DB 25; Length 383;
 Best Local Similarity 82.1%; Pred. No. 6,7e-75;
 Matches 307; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 QY 555 CCTGAGTATATGCGTGTAGCCCACTATGCTTGTCTGCACTGACGCTGCAACCAT 614
 DB 10 CATATGATATAGCGCTGTAGCCCACTATGCTTGTCTGCACTGACGCTGCAACCAT 69
 QY 615 GGTGCACTATGCGTGTAGCCCACTATGCGTGTAGCCCACTATGCTTGTCTGCACTGACGCTGCAACCAT 674
 DB 70 AGTGCACCTATGCGTGTAGCCCACTATGCGTGTAGCCCACTATGCTTGTCTGCACTGACGCTGCAACCAT 129
 QY 675 TGGCATAGTGTGCGCCATTTCTCATGAGTGGCCCAAGGTCCTTCAAGAGTGTGCGCC 734
 DB 130 TGGCACAAGTGTGCGCCATTTCTCATGAGTGGCCCAAGGTCCTTCAAGAGTGTGCGCC 189
 QY 735 TTTTCTGGAGAGAAAAGTCTATGCTGTGCAAGGCAAGGTCATATGCTTCTGCA 794
 DB 190 TTTTCTGGAGATTAAGGCTCTATGCTGTGCAAGGCAAGGTCATATGCTTCTGCA 249
 QY 795 AAGGGAGGCTATACCTTATAGCGGTTATCCGAACCGGCTGAGAGAGAGTGGAGAC 854
 DB 250 AAGGCAAGGCTATACCTTATAGCGGTTATCCGAACCGGCTGAGAGAGAGTGGAGAC 309
 QY 855 CCTCATGAGATATCTGAGACTGTGAGTGCAGGCTTTATCTGCGGCTTCTCGG 914
 DB 310 CCGAGAGGGGTCTGCTTCTATCTGTGAGTGCAGGCTTTATCTGCGGCTTCTCGG 369
 QY 915 CTCATATCATGCG 928
 DB 370 CTCATATCATGCG 383

RESULT 11
 ID ABX35215
 XX ABX35215 standard; cDNA; 437 BP.
 AC
 XX
 AC ABX35215;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #280.
 XX
 XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 KM muscle deposition; fat deposition; genome mapping; gene identification;
 KM gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 OS
 PN US2002137139-A1.
 PN
 PD 26-SEP-2002.
 PD
 PF 24-SEP-2001; 2001US-0960352.
 PF
 PR 12-JAN-1999; 99US-115707P.
 PR
 PR 11-JAN-2000; 2000US-0480902.
 XX
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 PI Byatt JC, Mathalagan N, Tao N, Warren WC;
 XX MPI; 2003-110599/10.
 DR
 XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX
 PS Claim 2; SEQ ID No 380; 245bp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are: (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.
 CC
 XX
 SQ Sequence 437 BP; 100 A; 102 C; 128 G; 104 T; 3 other;
 Query Match 8.3%; Score 264.4; DB 25; Length 437;
 Best Local Similarity 84.6%; Pred. No. 4.4e-74;

QY 1007 GGTAGACGAGCCCTTGATGAGAAAGTCCCTGAGCCCTAATCATGTTCCGCCAATGG 1066
 DB 241 AGTCGATGGGGCCCTGTGTACAGAGAGTCTCTGGGCCCCACATCGTTCTCCCAATGG 300
 QY 1067 TCCCGGCTTACTCTCATCATGAGTCCCAATTTGTACTGCTACAGATGTGAGAAACT 1126
 DB 301 TCTGGGCTGTACTCTGTCCAGGCCCAATCTGTACTGCTACAGAAATGTGAGAAAT 360
 QY 1127 GAATGACGACGAGCCCTTCCGCAAGCCCAATGTGACAGTCTCT 1174
 DB 361 GAGCAAGACCAAGACCTTCCCGAGGCCCAAGAGATGAACAGCTCT 408

RESULT 9
 ID ABX40203 standard; cDNA; 374 BP.
 AC ABX40203;
 DE 20-FEB-2003 (first entry)
 XX Bovine EST associated with lactation/muscle/fat deposition #5368.
 XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 XX muscle deposition; fat deposition; genome mapping; gene identification;
 XX gene analysis; cattle breeding.
 OS Bos Taurus.
 XX
 XX US2002137139-A1.
 XX 26-SEP-2002.
 XX 24-SEP-2001; 2001US-0960352.
 XX 12-JAN-1999; 99US-115707P.
 XX 11-JAN-2000; 2000US-0480902.
 XX
 XX (BYAT/) BYATT J C.
 XX (MATH/) MATHIALAGAN N.
 XX (TAON/) TAO N.
 XX (WARR/) WARREN W C.
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 DR WPI; 2003-110599/10.
 XX
 XX New nucleic acid associated with lactation, and muscle and fat
 XX deposition, useful for genome mapping, gene identification and
 XX analysis, cattle breeding, or for genetically improving cattle
 XX
 PS Claim 2; SEQ ID No 5368; 245bp; English.

The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 1512 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are; (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 1512 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern

CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 1512 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocId=20020137139.
 XX
 SQ Sequence 374 BP; 86 A; 102 C; 92 G; 94 T; 0 other;
 Query Match 8.4%; Score 269.2; DB 25; Length 374;
 Best Local Similarity 83.0%; Pred. No. 1.1e-75;
 Matches 307; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 555 CCTGATATATATGAGGCTGTAGCCCACTAGTCTTGTCTGCACTGACGTGCAACCAT 614
 DB 5 CATATGATTAAGGCTGTAGCCCACTAGTCTTGTCTGCACTGACGTGCAACCAT 64
 QY 615 GGTGACCACTATGAGCTTCACTGAGGAGCCCACTAGTCTGAGACAGCCGGAGTGG 674
 DB 65 AGTGCACCTTACGCTTTCAGTGAAGACCACTAGTCTGAGACAGCCGGAGTGG 124
 QY 675 TGGCATATGCTGAGCCCACTAGTCTGAGTGGCCCAAGGCTCTTACAGAGTGTGCC 734
 DB 125 TGGCAGAGCTGAGGAGTGTGAGATCTGTGGCCCAAGGCTCTTACAGAGTGTGCC 184
 QY 735 TTTTCTGGGAGAAAGAACTTATCTGTGTCAGGAGCAACCAAGTATATGCTTCTGACA 794
 DB 185 TTTTCTGGGAGTAAAGGCTTATCTGATCAGGAGTAAAGTATATCTTCTGACA 244
 QY 795 AAGGAGGCTATATACCTAGTAAAGCGGTATCCGAAGCGCTGAGAGAGAGTGGAGAC 854
 DB 245 AGGCGAGGCTTACACTCTCTTAAAGATTTATCCAAAGCACTGAGAGAGATTTGGAGC 304
 QY 855 CCTGATGAGATTAATCTGAGTCTGTGATGCGGCTTATCTGCGCTTCTCGG 914
 DB 305 CCTGATGAGGATCTGAGTCTGTGATGCGGCTTATCTGCGCTTCTCGG 364
 QY 915 CTTCATATCA 924
 DB 365 CTTCATATCA 374

RESULT 10
 ID ABX41681 standard; cDNA; 383 BP.
 AC ABX41681;
 DE 20-FEB-2003 (first entry)
 XX Bovine EST associated with lactation/muscle/fat deposition #6846.
 XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 XX muscle deposition; fat deposition; genome mapping; gene identification;
 XX gene analysis; cattle breeding.
 OS Bos Taurus.
 XX
 XX US2002137139-A1.
 XX 26-SEP-2002.
 XX 24-SEP-2001; 2001US-0960352.
 XX 12-JAN-1999; 99US-115707P.
 XX 11-JAN-2000; 2000US-0480902.
 XX
 XX (BYAT/) BYATT J C.
 XX (MATH/) MATHIALAGAN N.
 XX (TAON/) TAO N.
 XX (WARR/) WARREN W C.

CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule, and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridization between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.

CC Sequence 420 BP; 90 A; 114 C; 116 G; 100 T; 0 other;

Query Match 9.4%; Score 298.4; DB 25; Length 420;
 Best Local Similarity 81.9%; Pred. No. 4.6e-85;
 Matches 344; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 663 AGCCGGGATGCTGGCATAGCTGGCCCATTTCTCATCAGTGGCCCAAGGCTCTTACAGA 722
 DB 1 AGCCGGGAGCGGGTGGCAAGCTGCTCATTTGACATCTGTGGCCCAAGGCTCTTACAGA 60
 QY 723 GTGGATGCTGCTTTCTTGGGGAAGAAACCTATCTGCTCCAGGCGACCCAGGTATAT 782
 DB 61 GTGGATGCTGCTTTCTTGGGGAAGAAACCTATCTGCTCCAGGCGACCCAGGTATAT 120
 QY 783 GTCTTCTCTGACAAAGGAGGCTATACCTATAGAGGCTTATCCGAAGCGGCTGAGAA 842
 DB 121 ATCTTCTCTGACAAAGGAGGCTATACCTATAGAGGCTTATCCGAAGCGGCTGAGAA 180
 QY 843 GAAGTCGGGACCCCTCATGGAGATTATCTGAGACTGTGTGAGCGGCTTTATCTCCCT 902
 DB 181 GAATTTGGGAGCCCTCATGGGCTCTGCTCATTTGTGGATGACACCTTTACTCTGCT 240
 QY 903 GGGTCTTCTGCTCATATCATATCATGAGACGCGGCTGTGGTGTGACCTGAAGTCA 962
 DB 241 GGAATCTTCTGAGCTCATATCATATCATGAGACGCGGCTGTGGTGTGACCTGAAGTCA 300
 QY 963 GGAGCCCAAGCCACGTGACAGAGCTTCTTGGCCCATGAGAAAGGTAGACGAGGCTTG 1022
 DB 301 GGAGCTCAAGCCACGTGACAGAGCTTCTTGGCTCTCATACGAAGTCAATGGGCGCTTG 360
 QY 1023 TGTATGAAAAAGTCCCTTGGCCCTTAATCATATGTTCCGCCAATGATCCCGCTTGAACCT 1082
 DB 361 TGTATGAAAAAGTCCCTTGGCCCTTAATCATATGTTCCGCCAATGATGCGGCGCTTGAACCT 420

RESULT 8

ABX41875
 ID ABX41875 standard; cDNA; 408 BP.

XX ABX41875;

XX 20-FEB-2003 (first entry)

XX Bovine EST associated with lactation/muscle/fat deposition #7040.

XX Bovine; ss; EST: expressed sequence tag; lactation; LMPD;

XX muscle deposition; fat deposition; genome mapping; gene identification;

XX gene analysis; cattle breeding.

XX B08 Taurus.

PN US2002137139-A1.
 XX 26-SEP-2002.
 XX 24-SEP-2001; 2001US-0960352.
 XX 12-JAN-1999; 99US-115707P.
 XX 11-JAN-2000; 2000US-0480902.
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WAR/) WARREN W C.
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 DR WPI, 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle -
 XX Claim 2; SEQ ID No 7040; 245bp; English.

CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived
 CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are; (1) a transformed cell having a nucleic acid
 CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridization between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMPD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMPD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139.

XX Sequence 408 BP; 95 A; 109 C; 111 G; 93 T; 0 other;

Query Match 8.8%; Score 280; DB 25; Length 408;
 Best Local Similarity 80.4%; Pred. No. 3.9e-79;
 Matches 328; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 767 GGACACCCAGGTATATGTTCTTCTGACAAAGGAGGCTATATCCTAGTAAGCGGTTATCC 826
 DB 1 GGATACCCAGGTATATATCTTCTGACAAAGGAGGCTATATCCTAGTAAGCGGTTATCC 60
 QY 827 GAAGCGGCTGGAAGAAAGTGGGACCCCTCATATGAGGATTAATCTGAGACTCTGTGATCC 886
 DB 61 AAAGCAGCTGGAAGAAAGTGGGAGCCCTATATGAGGCTCTGCTTCAATCTGTGATCC 120
 QY 887 GGCCTTATCTGCTCCCTGAGGCTTCTTCTGAGCTCATATCATAGGAGACGCGGCTGTGATG 946
 DB 121 AGCTTTACTCTGCTGAGATCTTCTGAGCTCATATCATAGGAGACGCGGCTGTGATG 180
 QY 947 GCTGACCTGAAGTCAAGACCCCAAGCCAGTGAACAGAGCTTCTTGGCCCATGAGAA 1006
 DB 181 GCTGACCTGAAGTCAAGACCCCAAGCCAGTGAACAGAGCTTCTTGGCTCATATGAA 240

DB 1482 TTCCTCATATAAAGACAGATTGCTTCTTGCTTCACAGAGGGGCGCTTCGATGAG 1541
 QY 1292 TCTGGCTTGGCCCGCCACCTCCCGATTTCATATAAAGACAGATTGCTTCTTCATTGA 1351
 DB 1542 TCTGGCTTGGCCCGCCACCTCCCGATTTCATATAAAGACAGATTGCTTCTTCATTGA 1601
 QY 1352 ATCAA 1356
 DB 1602 AAAAA 1606

RESULT 3
 AB283572
 ID AB283572 standard; cDNA; 690 BP.
 AC AB283572;
 XX 14-MAY-2003 (first entry)
 DT Toxicologically relevant human nucleotide sequence #731.
 XX Toxicologically relevant gene; toxicological response; gene; ss.
 OS Homo sapiens.
 XX MO2003016500-A2.
 XX 27-FEB-2003.
 PD 16-AUG-2002; 2002WO-US26514.
 PF 16-AUG-2001; 2001US-313080P.
 PR (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
 PA Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeisler K;
 PI Alen P;
 XX MPI; 2003-268322/26.
 DR Determining a toxicological response to an agent, useful for screening
 PT of drugs, comprises comparing the expression profile of one or more
 PT human toxic response genes to a reference gene expression profile
 PT indicative of toxicity -
 PS Claim 1; Page 241; 455pp; English.
 XX The present invention describes a method (M1) for determining a
 CC toxicological response to an agent, which comprises comparing the
 CC expression profile of one or more human toxic response genes to a
 CC reference gene expression profile indicative of toxicity, and so
 CC determining the presence of a toxic response to the agent. Also
 CC described: (1) an array comprising one or more polynucleotides selected
 CC from the genes corresponding to the partial sequences given in AB282842
 CC to AB284764, or their fragments of at least 20 nucleotides, or
 CC homologues; and (2) determining if a gene putatively identified to be a
 CC toxic response gene plays a role on toxic response pathways by
 CC determining the expression profile of the gene after exposure of cells
 CC or a human subject to a known toxic pharmaceutical or industrial agent,
 CC comprising: (a) exposing cells to an agent or isolating cells from a
 CC human subject who was exposed to an agent; (b) obtaining the test gene
 CC expression profile for a putatively identified toxic response gene after
 CC exposure to a known toxic pharmaceutical or industrial agent; and
 CC (c) comparing the test profile to the expression profile of a gene with
 CC a similar function or comparing the test profile to the expression
 CC profile of that gene after exposure to other known toxic compounds. The
 CC methods are useful for predicting and determining toxicological responses
 CC on a cellular, organ or system level. The arrays comprising the human
 CC genes are useful for toxicological screening of drugs, pharmaceutical
 CC compounds and chemicals.
 XX Sequence 690 BP; 147 A; 181 C; 196 G; 164 T; 2 other;

Query Match 16.0%; Score 509.4; DB 25; Length 690;
 Best Local Similarity 99.8%; Pred. No. 5.5e-153;
 Matches 510; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 630 TTCAAGTGGACCCCACTACTAGGCGTCTGACACACAGCCGGGATGCTGGCATAGCTGGCC 689
 DB 72 TTCAGTGGAGCCCACTACTAGGCGTCTGACACACAGCCGGGATGCTGGCATAGCTGGCC 131
 QY 690 ATTGCTCATCAATGAGCCCGGAGGCTCTTACGACAGTGGATGCTGCTTTCTTGGAGAA 749
 DB 132 ATTGCTCATCAATGAGCCCGGAGGCTCTTACGACAGTGGATGCTGCTTTCTTGGAGAA 191
 QY 750 AAACCTATCTGATCTGACAGGACCCAGGATATGCTTCTTCTGACAAAGGAGGCTTATCC 809
 DB 192 AAACCTATCTGATCTGACAGGACCCAGGATATGCTTCTTCTGACAAAGGAGGCTTATCC 251
 QY 810 CTAGTAAAGCGGTTATCCGAAAGCGGCTGAGAAAGAAAGTGGAGACCCCTGATGGATATC 869
 DB 252 CTAGTAAAGCGGTTATCCGAAAGCGGCTGAGAAAGAAAGTGGAGACCCCTGATGGATATC 311
 QY 870 CTGACTCTGTGATGAGGCGGCTTTATCTGCGCTCTTCTGCGTCCATATCATGAGCA 929
 DB 312 CTGACTCTGTGATGAGGCGGCTTTATCTGCGCTCTTCTGCGTCCATATCATGAGCA 371
 QY 930 GGACGCGGCTGTGATGAGGCTGAGACCTGAAGTCAAGAGCCCAAGCCAGTGAAGAGCTT 989
 DB 372 GGACGCGGCTGTGATGAGGCTGAGACCTGAAGTCAAGAGCCCAAGCCAGTGAAGAGCTT 431
 QY 990 CCTTGGCCCATGAGAAAGTACGAGAGCTTGTGTATGAAAGTCCCTTGGCCCTTAAC 1049
 DB 432 CCTTGGCCCATGAGAAAGTACGAGAGCTTGTGTATGAAAGTCCCTTGGCCCTTAAC 491
 QY 1050 TCATGTTCCGCCAATGATGCTCCGCTTTTATCTCATATCATATGATGCTCCCATTTGATCTCTAC 1109
 DB 492 TCATGTTCCGCCAATGATGCTCCGCTTTTATCTCATATCATATGATGCTCCCATTTGATCTCTAC 551
 QY 1110 AGTGAATGAGAACTGAATGACAGCAAG 1140
 DB 552 AGTGAATGAGAACTGAATGACAGCAAG 582

RESULT 4
 ABT08987
 ID ABT08987 standard; DNA; 631 BP.
 AC ABT08987;
 XX 05-DEC-2002 (first entry)
 DT Phase-1 Rat CT gene SEQ ID No 75.
 XX Phase-1 Rat CT gene SEQ ID No 75.
 DE Rat; toxicity study; rat toxic response gene; toxicological response;
 KW drug development; phase-1 rat CT gene; ds.
 XX Rattus sp.
 OS MO200266682-A2.
 PN 29-AUG-2002.
 XX 29-JUN-2002; 2002WO-US02935.
 PD 29-JUN-2001; 2001US-264933P.
 PF 29-JUN-2001; 2001US-264933P.
 PR 26-JUL-2001; 2001US-308161P.
 XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
 PA Farris G, Hicken SH, Farr SB;
 PI MPI; 2002-674961/72.
 DR Evaluating the toxicity of an agent, useful in drug development or in
 XX determining toxicological responses to a new drug, by determining the

PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -

XX Claim 1; Page 244; 327bp; English.

CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by them are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a
CC cancer, immunopathology or neuropathology.

XX Sequence 1571 BP; 344 A; 431 C; 434 G; 362 T; 0 other;

Query Match 35.4%; Score 1126.8; DB 22; Length 1571;
Best Local Similarity 86.3%; Pred. No. 0;

Matches 1351; Conservative 0; Mismatches 2; Indels 213; Gaps 1;

QY 4 TGCAGCTCAGCATGCTAGAGGTACTGAGAGCACCCTGTCATGAGGTTGTGAGCTAT 63
DB 1 TGCAGCTCAGCATGCTAGAGGTACTGAGAGCACCCTGTCATGAGGTTGTGAGCTAT 60
QY 64 GCTGCTCTGCGCATTTGCCACCCCTCTTCTCCGACTAGTCCCATGGAATGTTGCTG 123
DB 61 GCTGCTCTGCGCATTTGCCACCCCTCTTCTCCGACTAGTCCCATGGAATGTTGCTG 120
QY 124 AAGGCGAGACCAAGCCAGACCCGAGTGAAGTGAAGCTGCTGATGCTGAGGCTTGG 183
DB 121 AAGGCGAGACCAAGCCAGACCCGAGTGAAGTGAAGCTGCTGATGCTGAGGCTTGG 180
QY 184 ATGCTACCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243
DB 181 ATGCTACCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 244 AGAGTCAAAATGGGACCGGAGTTATCTCAAGAGATGGAATTTCCCAAGCCCTG 303
DB 241 AGAGTCAAAATGGGACCGGAGTTATCTCAAGAGATGGAATTTCCCAAGCCCTG 300
QY 304 TGGATGCTGATCCGTCAAAGTCACAAAGTCCTTTTCAAGAGGAGGAGCAAAAGTCT 363
DB 301 TGGATGCTGATCCGTCAAAGTCACAAAGTCCTTTTCAAGAGGAGGAGCAAAAGTCT 360
QY 364 GGGTATACCTCCTGAAAAGAGAGAGAGATACCAAGTTGCTCAAGATGATGATTTT 423
DB 361 GGGTATACCTCCTGAAAAGAGAGAGATACCAAGTTGCTCAAGATGATGATTTT 420
QY 424 CTGGAATCCCATCCCACTGAGATGAGCTGTGGAATGTCACCGTGAAGATGTCAGCTG 483
DB 421 CTGGAATCCCATCCCACTGAGATGAGCTGTGGAATGTCACCGTGAAGATGTCAGCTG 480
QY 484 AAGGCGTCCCTCTTCCCA----- 502
DB 481 AAGGCGTCCCTCTTCCCAAGGAGTGAACCGGAGTGTGTTGAGGACTTGGGTAACGGAACCA 540
QY 503 ----- 502
DB 541 TGAAGAGAGGTTCTGAGCAGCTGTTGGAACTGCTCTGCTGAGATGAGCTGGGCC 600
QY 503 ----- 502
DB 601 GCTACTACTGCTTCAGAGGTAAACCAATCTCTGCGCTTCGACCTGTGAGGAGAGAGTGC 660

QY 503 -----AGGCCATG 510
DB 661 CTCACAGTACCCCGGAGATGTCAGACTACTTCAATGCTCCCTGCGACAGAGGCTATG 720
QY 511 GACACAGAAATGGAGCTGGCCATGGAACAGTACCAACCATGAGCTCTGATATATGCGCT 570
DB 721 GACACAGAAATGGAGCTGGCCATGGAACAGTACCAACCATGAGCTCTGATATATGCGCT 780
QY 571 GATGCCACATATAGCTTGTGTGCACTGACATGACATGACATGACATGACATGACATG 630
DB 781 GATGCCACATATAGCTTGTGTGCACTGACATGACATGACATGACATGACATGACATG 840
QY 631 TCAGTGGAGCCACTACTGAGCTGAGACACAGCCGAGATGAGTGGATAGCTGAGCCCA 690
DB 841 TCAGTGGAGCCACTACTGAGCTGAGACACAGCCGAGATGAGTGGATAGCTGAGCCCA 900
QY 691 TTGCTCATGATGAGCCGAGGCTCTTCAAGATGATGCTGCTTCTTCTGGAAGAAA 750
DB 901 TTGCTCATGATGAGCCGAGGCTCTTCAAGATGATGCTGCTTCTTCTGGAAGAAA 960
QY 751 AACTATATGATGAGCCGAGGCTCTTCAAGATGATGCTGCTTCTTCTGGAAGAAA 810
DB 961 AACTATATGATGAGCCGAGGCTCTTCAAGATGATGCTGCTTCTTCTGGAAGAAA 1020
QY 811 TAGTAAAGCGTTATCCGAGAGCGGCTGAGAGAGAGTGGAGACCCCTCATGAGATATCC 870
DB 1021 TAGTAAAGCGTTATCCGAGAGCGGCTGAGAGAGAGTGGAGACCCCTCATGAGATATCC 1080
QY 871 TGGACTCTGATGATGAGGCTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
DB 1081 TGGACTCTGATGATGAGGCTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 931 GACGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 990
DB 1141 GACGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 991 CTTGAGCCCATGAGAGATGAGAGCGAGCTTGTATGAGAAAAGTCCCTTGGCCCTACT 1050
DB 1201 CTTGAGCCCATGAGAGATGAGAGCGAGCTTGTATGAGAAAAGTCCCTTGGCCCTACT 1260
QY 1051 CATGTTCCGCAATGATGCTCCGCTTGTATCTGATCATGATGATGATGATGATGATGAT 1110
DB 1261 CATGTTCCGCAATGATGCTCCGCTTGTATCTGATCATGATGATGATGATGATGATGAT 1320
QY 1111 GTGATGATGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170
DB 1321 GTGATGATGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1171 TCTGAGGCTGCACTGATGAGGAGCTTCTGACATGATGATGATGATGATGATGATGAT 1230
DB 1381 TCTGAGGCTGCACTGATGAGGAGCTTCTGACATGATGATGATGATGATGATGATGAT 1440
QY 1231 GTTCTCATATATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290
DB 1441 GTTCTCATATATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1291 GTTCTGAGCTGAGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350
DB 1501 GTTCTGAGCTGAGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1351 AATCAA 1356
DB 1561 AAAAAA 1566

RESULT 2

AAH98592
ID AAH98592 standard; cDNA; 1610 BP.

AAH98592;

12-OCT-2001 (first entry)

